For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex libris
Universitatis
Albertaensis

QUAECUMQUE VERA

Digitized by the Internet Archive
in 2021 with funding from
University of Alberta Libraries

https://archive.org/details/Reid1972

THE UNIVERSITY OF ALBERTA

THE AURAL DISCRIMINATION OF BASIC SENTENCE-TYPES

by

© JAMES RAYMOND REID

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA

FALL, 1972

Thesis
72F-188

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "The Aural Discrimination of Basic Sentence-Types", submitted by James Raymond Reid in partial fulfilment of the requirements for the degree of Master of Science.

ABSTRACT

An experiment was performed, extending the Baker, Prideaux, & Derwing ('BP&D') conjunctive concept-formation study (in press) in its assessment of the relative subjective prominence of syntactic features. A set of 128 sentences, systematically varied in Voice (active or passive), Mood (declarative or interrogative), and Modality (affirmative or negative), as well as tense and lexical content, was presented aurally, in one of eight random orders, to each of 32 male and 32 female undergraduate Ss. The task was to infer, from differential correct/wrong reinforcement of yes/no judgments, a target sentence-type characterized by one of the eight possible combinations of Voice, Mood, and Modality features, and to respond 'yes' to all and only such sentences.

An analysis of variance showed a significant difference among mean yes-response counts, with the bulk of yes-responses to target-type sentences or to Voice-opposites thereof; these means differed significantly from each other, and from all other error-type means ($p < 0.01$). Yes-response proportions, as a function of target-type and error-type, correlated highly with the corresponding BP&D values ($r = 0.90$). Since 20 of the present 64 Ss had failed to form a complete three-way conjunctive target

concept by the 128th trial, the distribution of means of last trials in which Voice, Mood, or Modality yes-errors were committed was not systematic; however, omission of these _S_s from the analysis yielded results commensurate with those of BP&D.

The present experiment was judged to have replicated BP&D in all essentials, and to have provided their conclusions with a broader base of support. The high 'failure' rate, and a comparison of the median 'trials-to-criterion' with those of previous studies suggested that the aural task was, in general, more difficult, yet entirely feasible, provided _S_s were not limited by a fixed number of trials. The relative obviousness of Mood and Modality features confirmed the traditional assumption that negation and interrogation are the 'marked' members of these dichotomies. _S_s' initial unawareness of the active/passive distinction indicated, however, that neither alternate is 'marked' as such, despite the restricted occurrence of passives. The conscious production and apprehension of passive focus seems to depend on extrasentential variables, which current formal grammar theory, as a matter of principle, excludes from consideration.

ACKNOWLEDGMENTS

This work would never have come to fruition without the generous help of a great many people, of whom I can thank only a few expressly.

To my wife Lorraine, for her unfailing patience and encouragement, this thesis is lovingly dedicated.

I am most indebted to Dr. Wm. J. Baker, my supervisor, the originator of the experiment reported herein. His ideas so permeate every aspect of this thesis that it is impossible to acknowledge them all individually. I thank him for these, and for his attentive guidance throughout the past two years. I am also grateful to Drs. Gary D. Prideaux and Jeff R. Sampson, the other members of my examining committee, for their helpful comments and suggestions.

I thank my instructors and my fellow students, present and past, for whatever they may have contributed toward this work, especially John Gray, who executed Fig. 1 on a CalComp plotter.

My two years of Master's study were supported by the University of Alberta and the Department of Linguistics, through graduate assistantships and intersession bursaries, and by the American Council of Learned Societies.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER ONE

INTRODUCTION

Background to the Problem

Performative Measures of Sentential Complexity

A great deal of attention has been devoted in the past decade to the psychological relevance of linguistic 'deep structure' and grammatical transformations. It has been argued that the production or understanding of natural language, since it is a case of 'making infinite use of finite means', cannot be adequately modeled except by a formal system whose fundamental component is a generative-transformational grammar (Miller & Chomsky, 1963). On this assumption, latter-day linguists have in the main concerned themselves with the formalization, on the generative-transformational model, of the mental grammar 'internalized' by every speaker of a language; grammarians claim thereby to characterize abstractly, and in complete isolation from linguistic performance, 'the ideal speaker-hearer's intrinsic competence' (Chomsky, 1959, pp. 56-7; 1965, pp. 4, 8, & passim).

As early as 1959, Chomsky was suggesting that "it should be possible to derive from a properly formulated grammar a statement of the integrative processes and generalized patterns imposed on the specific acts that constitute an utterance" (1959, p. 56). To substantiate this putative relation between mental grammar and the generative-transformational model thereof, Miller & Chomsky later proposed that:

> The psychological plausibility of a transformational model of the language user would be strengthened, of course, if it could be shown that our performance on tasks requiring an appreciation of the structure of transformed sentences is *some function of the nature, number, and complexity of the grammatical transformations involved* (1963, p. 481, emph. added).

This seemed a safe prediction at the time, for some encouraging experimental results already obtained by Miller and others supported the notion that "kernel sentences play a central role, not only linguistically, but psychologically as well" (*ibid*., p. 483).

Since then, many cognitive psychologists and psycholinguists have been attempting to determine the tenability of what is variously known as the Derivational Theory of Complexity (DTC; Fodor & Garrett, 1967), or the Correspondence Hypothesis (CH; Hayes, 1970). Both these notions propose a positive correlation between the psycholinguistic complexity of a sentence and the number of transformational operations required to specify its surface

structure in a formal grammar. The CH, furthermore, which appears to have motivated much of the earlier work, had a homuncular cast to it, implying a step-by-step correspondence between the logically-sequenced derivational procedures of the grammar and the temporally-sequenced recoding processes of the language-user. This simple-minded interpretation regarded the grammar as the whole performance model in itself, in flagrant disregard of many admonitions to the contrary (e.g., Chomsky, 1957, p. 48; Miller & Chomsky, 1963, p. 421). A third notion, also subjected to some experimental scrutiny during the same period, was the performance-oriented 'depth hypothesis' (Yngve, 1960). It, too, sought to correlate the perceptual complexity of sentences with their derivational complexity, but disavowed the relevance of transformations: surface structure was described entirely in terms of constituent-structure rules, and complexity was viewed as a function of 'unresolved constituents', rather than the number of rules *per se*.

The relevant work has been reviewed by a number of writers, including Fodor & Garrett (1966), Ervin-Tripp & Slobin (1966), Bever (1968), Schlesinger (1968), Goldman-Eisler & Cohen (1970), Fillenbaum (1971), and Gough (1971). The general consensus seems to be that a decade of experimentation has yielded nothing but conflicting results, none of them entirely convincing. The earlier findings, however, while perhaps not the unqualified "series of most impressive successes" Fodor & Garrett eulogize (1966,

p. 143), did appear to provide rather solid support for the CH. It was shown that:

1. Sentences whose formal derivation requires more transformations are psychologically more difficult to handle than transformationally simpler ones, in that:
   (a) they require more time to read, understand, produce, evaluate, or recall (Miller, 1962; McMahon, 1963; Mehler, 1963; Miller & McKean, 1964; Gough, 1965, 1966), or
   (b) they place a greater burden on the constant, limited capacity of short-term memory (Savin & Perchonock, 1965);
2. Erroneously recalled or recognized sentences tend to be transformationally simpler than their correct counterparts, lending credence to the 'kernel plus syntactic footnote' view of recoding (Miller, 1962; Mehler, 1963);
3. Differences among sentences in the lengths of their transformational 'histories' are reflected in behavioral measures of their similarity or confusibility (Clifton, Kurcz, & Jenkins, 1965).

But these initial studies were limited in scope, with respect to both the syntactic complexity of materials employed and the behavioral correlates being measured. Subsequent work incorporated a greater variety of independent and dependent variables, seeking to confirm and

extend the first findings; support of the CH or DTC in these later experiments turned out rather equivocal, however, compared to the earlier results. At the same time, the extensive revisions in grammar theory proposed by Fodor, Katz, Postal, Chomsky, and others (see Clifton & Odom, 1966, pp. 5-8) began to gain acceptance among psycholinguists; this gave rise to a divergence of theoretical viewpoints and specific experimental hypotheses, and led to the eventual tacit abandonment of the simplistic CH. Among these inconclusive results were the following:

1. Johnson (1965) claimed he was inducing the subjective phrase-structure of simple sentences, but, since his materials included transformationally embedded constituents of various kinds and lengths, his failure to find differential learning-rates in correspondence with sentence-types cannot be interpreted as evidence either for or against the CH and the DTC.
2. Slobin (1966) obtained the DTC-predicted ordering relations among sentences of varied transformational complexity, but only where the grammatical subjects and objects of such sentences could assume either of the semantic roles of 'actor' or 'acted-upon'; 'non-reversibility' made transformationally more complex passives performatively indistinguishable from simpler sentences.
3. Clifton & Odom (1966) had nothing to say about either CH or DTC, but found "substantial evidence for the

existence of some parallel between the linguistic description of a language and the reactions of a language user to his language" (p. 32). Their attempts to quantify perceived similarity among syntactically-related sentences, and their re-analysis of Mehler's (1963) data along the same lines, were generally in accord with a prediction model inferred from the revised grammar theory of Katz & Postal (1964). Some distortion of the expectations of the older 'transformational distance' view was involved, but considerable congruence remained. Thus, Clifton & Odom's findings did not as such contradict the DTC, but quantified and refined its predictions, in accordance with the new 'competence' grammar.

4. Fodor & Garrett (1966) reported a number of unpublished failures to corroborate earlier findings, on which basis they denied the psychological reality of the "operations whereby grammars generate structural descriptions" (p. 152). However, they held the psychological reality of grammar itself to be, because of the 'strong internal evidence' in its favor, unassailable, and they argued that the linkage between competence and performance must be 'more abstract' than the CH had implied. In Fodor & Garrett (1967) this more abstract relationship turned out to be that the psychological complexity of a sentence depends not (only) on the transformational distance between deep

and surface structure, but also, since "transformations tend to destroy structure" (p. 290), on the extent to which surface clues for the listener's correct reconstruction of deep structure have been transformationally deleted. How one might extricate these factors one from another, they did not say, and the experimental evidence they adduced in support of this paradoxical position was unconvincing. For, while their findings more or less conformed to some scale of 'elaboratedness of structure' in sentences, the performative differences induced through added adjective-embedding tended in one direction with visual presentation and in the other with auditory, with neither difference statistically significant. Their 'more abstract relation' thus remained largely speculative, and their claim that "the drastic performance decrement predicted by DTC is not forthcoming" (1967, p. 294) was but a futile attempt to prove the Null Hypothesis.

5. Perfetti (1969), attempting to discredit the Yngve depth hypothesis, attributed the poorer recall of sentences not to greater structural mean depth, but to greater proportions of lexical content words <u>vis a vis</u> function words. He noted, however, that this 'lexical density' factor directly reflected the number of (transformationally embedded) 'assertions' which underlie a sentence (see Chomsky, 1965, pp. 128ff;.

Jacobs & Rosenbaum, 1968, pp. 199ff.). No firmer conclusion could then be drawn than that "underlying structures are relevant in retention" (p. 723), for this finding seemed to support the transformation-count metric of sentential complexity as much as lexical density.

A number of other workers either could find no significant difference at all in dependent measures presumed to be related to transformational complexity, or found evidence which contradicted the CH and DTC. Matthews (1968) was unable to replicate Savin & Perchonock's (1965) finding of systematically differentiated recalls of extraneous words along with sentences of varied complexity. Martin & Roberts (1966) found Yngve's surface-structure-oriented 'mean depth' a better predictor of ease of recall than transformational sentence-type, this latter factor appearing to be of but marginal import: not only did they observe a partial reversal of DTC predictions - the simplest sentences, kernels, were worst recalled - but their subjects' (Ss') errors tended to simplify mean depth and left transformational complexity unaltered, contrary to what Mehler's (1963) results would have predicted. Rohrman (1968) found that 'subject nominalizations' (subordinate sentences, embedded as verb complements) tended consistently to facilitate recall, contrary to DTC predictions. And Slobin (1968), extending his 1966 work with passives, found that the transformationally more complex truncates tended to

be recalled verbatim by both children and adults, whereas full passives tended to be 'simplified'.

Some Theoretical Objections

The most glaring inadequacy in the studies reviewed above is the bidirectional confounding of sentence-length with transformational history. Since transformations (T-rules) can add, delete, or substitute constituents, derivationally 'complex' sentences can be longer than, or shorter than, or even the same length as, 'simple' ones. But then, why should derivational complexity be quantified solely in terms of T-rules? Chomsky claimed, very early on, that:

> . . . the optional rules of the grammar [could be viewed] as the selective mechanisms involved in the production of a particular utterance . . . The speaker's task is to select a particular compatible set of optional rules . . . The listener (or reader) must determine, from an exhibited utterance, what optional rules were chosen in the construction of the utterance" (1959, pp. 56-7).

No distinction was made then among phrase-structure (PS-) rules, phonological (P-) rules, and transformations, all of which embodied some degree of optionality. The derivational histories of sentences could thus differ as much on account of the number of PS-rules or P-rules applied as on account of the number of T-rules, and the singling out of optional transformations as the key to performative complexity

(Miller & Chomsky, 1963; see p. 2, above) now seems to have been somewhat arbitrary.

Optional transformations, furthermore, were an ill-starred feature of early Chomskyan grammars. Since 1957, generative-transformational theory has undergone considerable revision, and quite a different view is now taken of optionality in language-description (see Katz & Postal, 1964; Chomsky, 1965). The difference is a subtle one: T-rules which once optionally transformed structures underlying one sentence-type into those of another are now obligatory, specifying the ordering and dominance relationships through which certain abstract semantic sentence-type features are made manifest. Such transformations do not change meaning as such, but only specify the conventional language-specific syntactic shapes for sentence-meaning features already inherent in deep structure. The only transformations which remain optional in 'post-*Aspects*' grammars (see Chomsky, 1965) are those which characterize purely stylistic variation among sentences having the 'same meaning'.

Semantics, which was virtually ignored in early generative-transformational theory and in most of the work reviewed here, may turn out to be a much greater factor in linguistic processing than linguistics has heretofore allowed. There was, for example, an obvious confounding of lexical meaning with syntax in Slobin's psychologically more

complex 'reversible' passives (1966), and Miller & McKean were convinced, very early in the decade, that

> . . . These [syntactic] operations are quite automatic and involuntary, as are most of the information processes that psychologists have postulated to occur in perception or in motor skills. Any deliberate and explicit control that a person can exercise over these linguistic operations is principally, and perhaps exclusively, at the semantic level (1964, p. 307).

Both CH and DTC depended on formal grammar to order linguistic constructions with respect to their syntactic complexity. But grammar has always done so in a manner inconsistent with what every speaker 'knows'. Whereas negation and interrogation are both semantically and syntactically different from simple assertion, passivization has long been known to involve considerable syntactic permutation and addition, with little or no concomitant difference in sentence-meaning (see Chomsky, 1957, pp. 94, 100-1; Katz & Postal, 1964, pp. 72-3). In their single-minded pursuit of theoretical simplicity, however, grammarians have tended to treat these and other syntactic 'processes' as qualitatively and quantitatively no different, whether as optional transformations (Chomsky, 1957), or as optional 'universal morphemes' in deep structure (Katz & Postal, 1964, pp. 118-9; Chomsky, 1965, pp. 132ff.). Most psycholinguists, furthermore, seem to have taken the same view: performatively, as well as formally, one transformation was as good as another (see

Clifton & Odom, 1966). Small wonder, then, that mental grammars could not be found to correspond consistently with theoretical grammar!

Methodological Problems

Many of the studies cited above are also open to direct criticism on methodological grounds, in that they often involved the use of what Fillenbaum (1970) calls "memorial" techniques, *e.g.*, immediate recall, free recall, or recognition. What tended to be overlooked was that the data from such experiments are subject to at least two influences:

1. the 'true' treatment effect(s), if any, attributable to the independent variable(s) under the experimenter's (*E*'s) control, and
2. the effects of memory, retrieval, and other processes not controlled in the studies in question, and therefore inextricably confounded with the observed treatment effect(s).

It is a commonplace that structured verbal materials are easier to process and remember than unstructured; this is generally attributed to 'chunking', with (7 ± 2) 'chunks' available for immediate recall in most instances (Miller & Selfridge, 1950; Miller, 1956; Miller & Isard, 1963; Marks & Miller, 1964). What remains not too well-known, however, is how variations in syntactic structure

affect the constitution and cohesion of chunks: quite long 'right-branching' multiply-embedded structures appear to be easier to understand, contrary to expectations, than relatively short 'center-branching' (self-)embeddings (Miller, 1962; Miller & Isard, 1964). Nor has it yet been reliably established whether chunks correspond in any systematic way to syntactic constituents (Fodor & Bever, 1965; Garrett, Bever, & Fodor, 1966; Bever, Lackner, & Kirk, 1969; but see Reber & Anderson, 1970). Techniques involving undue reliance on memory are therefore to be eschewed, in favor of more direct methods of assessment.

Goldman-Eisler & Cohen (1970) have hypothesized that, in view of the enormous disproportionality of occurrence - and hence, of familiarity - among even common sentence-types, behavioral differentials may not be a valid measure of syntactic complexity at all. Familiarity and complexity are so highly (and negatively) correlated, it is suggested, that the observed facts can be accounted for by the one as easily as by the other. This claim is not entirely borne out, however, for the distribution observed by these authors would predict little or no difference among non-kernel sentences, and a vast difference between them and kernels; such has not generally been the case in the work reported so far. Indeed, the only general tendency there seems to be in the data is towards unpredictability, and that alone is a strong indication of uncontrolled variables at work. Thus, while differential familiarity in structured verbal

materials may exert an influence of its own on dependent measures, this factor has been just another overlooked source of variation, not a reason for abandoning psychosyntactic experimentation.

But it is not in independent effects that the familiarity factor is likely to be a source of bias, for a complete factorial crossing of sentence-types with the other independent variables in an experiment would tend to counterbalance such effects and make them constant (see Mehler, 1963; Clifton & Odom, 1966). It is rather in its interactions with uncontrolled variables such as memory processes that it threatens the validity of data. Obviously, if syntactic complexity is highly correlated with familiarity, the ease with which any type of sentential stimulus is processed or remembered (or forgotten!) cannot be attributed to either factor with certainty. But if, as in the natural context of language use, there is no necessity to recall verbatim the linguistic stimulus, nor any constraint tc process it quickly, familiarity effects can be expected to be relatively trivial. Thus, the implementation of minimally memory-dependent assessment techniques, already espoused for other reasons, procures a further advantage.

Foundational Inadequacies

Quite apart from the theoretical and methodological difficulties outlined above, considerations of an even more fundamental nature lead one to wonder how hypotheses such as the CH or DTC were ever seriously entertained in the first place. A transformational grammar is, after all, only an abstract, formal rewriting system which specifies a denumerably infinite set of (by stipulation) well-formed symbol-structures. But, on the basis of the well-known 'fact' that these symbol-structures are isomorphic with the structural descriptions of well-formed sentences, functional equivalence, if not isomorphism, has been postulated to hold between our mental grammar and the abstract formal device. Such a hypothesis now appears simplistic, for it arises from an all too easy acceptance by psychologists of the 'competence' claims of transformational linguists, and it embodies a myopic view, on the part of both, of what Derwing calls "the crippling _inferential gap_ between the linguist's grammar and the bulk of the observable facts of linguistic behavior" (in press, pp. 274-5).

The stronger hypothesis - the CH - is, of course, patently indefensible, for it requires one to believe, among other things, that speaker-hearers actually perform some mental analog of, say, the 'affix-shift' transformation. This is the T-rule which formalizes the fact that, in English verb-phrases, the tense-marker is always suffixed to

the first verbal element, and the '-*en*' and '-*ing*' particles always postposed to the morpheme following perfective *have* and progressive *be*, respectively. One might wonder in what sense such a T-rule expresses the formal syntactic relations between two classes of sentences. The answer, of course, is that it does not: the English facts in question are formulated as a T-rule for no other reason than that, in order to "duplicate the effect of [affix-shift] without going beyond the bounds of a system $\{\Sigma, F\}$ of phrase structure, it would be necessary to give a fairly complex statement." (Chomsky, 1957, p. 41)

There can be no quarrel with a desire on the part of grammarians to achieve theoretical elegance through simplicity; to impute the same goal to language-learners, however, seems somewhat visionary. For, as Watt has suggested, "the facts that the [linguists' grammar] is a grammar, and that the linguistic faculty appears to include a grammar, never strictly entailed that the two grammars were essentially identical." Grammar theory may well value overall economy, and thus make all possible significant generalizations, but there is no reason whatever to believe that communication-oriented language-users are similarly motivated (Watt, 1970b, pp. 187ff.). They, after all, bring to every speech act a whole world of presuppositions which the formal automaton cannot begin to formalize adequately, despite current efforts to make it do so.

Even the weaker hypothesis, DTC, does not stand up under close scrutiny, for it is founded on the assumption that formal transformational grammar correctly represents the speaker-hearer's 'tacit knowledge', and mirrors the compositional manner in which a sentence is produced or understood (Katz & Fodor, 1963, p. 482). Now, it is not unreasonable to suppose that our mental grammar ('competence', 'tacit knowledge', or whatever), through the mediation of certain little-known psychological input/output procedures, could give rise to a denumerably infinite set of utterances, some clearly well-formed, some clearly not. And it may well be that the symbol-structures enumerated by formal grammar map perfectly onto the structural descriptions ascribed to those well-formed utterances. But that should come as no great surprise, for it is exactly what grammarians invent grammars for - to reproduce in mindless fashion their enlightened linguistic intuitions about the real or potential well-formed sentences of natural language L. However, since well-formed sentences are ideal fictions abstracted from the set of all possible utterances in L, what a 'grammar of L' describes is not really L at all, but only a restricted set of imaginary structures which are inferrable from an equally unsubstantial error-free subset of L.

Such theoretical 'prediction' is essentially vacuous, as Stuart has pointed out, for "there is no provision for

claiming that the relational structure expressed in the description has any further ontological status than that given by the rational mapping relation" (1969, p. 396). It becomes doubly trivial when linguistic theory insists that the ultimate test of a grammar is that which gave rise to it in the first place, _i.e._, the informed intuitions of linguistic analysts about the ideal fictions their grammar was designed to generate (see Chomsky, 1965, pp. 18-21; Leech, 1968, p. 90). Generative-transformational theory thus rests on a foundation which is not only fraught with "vicious circularity" (see Harris, 1970, pp. 35-43), but also resolutely cut of contact with reality, for its principal source of validation is only "a matter of goodness-of-fit with respect to the primary data, and not a matter of testing hypotheses about empirical states of affairs that underlie these data" (Stuart, 1969, p. 397). Only those possessed of inordinate belief in the "oracular infallibility of intuition" (Leech, 1968, p. 99) - or, as some call it, 'strong internal evidence' (Fodor & Garrett, 1966, p. 152) - could seriously attribute empirical predictiveness to formal linguistic descriptions on such an ephemeral basis. And only those willing to believe _them_ would attempt to measure the functional equivalence of formal grammar and mental grammar without considering whether such relatedness can even be said to exist.

The Problem

'Syntactic' versus Semantic Processing

It seems fairly clear that most attempts to determine *how* language-users map structural descriptions onto acoustic waveforms (and vice versa) have been only partly successful. A number of reasons therefor have been proposed here, ranging from the methodological to the epistemological: in general, the experimental question has been one that cannot logically be asked, and the procedures used to answer it have often been inappropriate to the study of natural language. And yet, the experiments have not been totally uninformative, or else interest in the matter would not have persisted for so long. The fact remains that, when people are subjected to systematically differentiated linguistic stimuli, they respond in systematically differentiated ways, though perhaps not in accord with theoretical expectations. It is perhaps time to revise those expectations, or as L. J. Cohen has suggested, to reconsider the whole question of whether speaker-hearers map waveforms onto structural descriptions at all (in Lyons & Wales, 1966, pp. 163-73).

Chomsky has claimed repeatedly that to understand a sentence *is* to infer the structural description which a correct grammar would assign to it (1959, pp. 56-7; 1964, p. 10; 1965, pp.4-5). Every attempt to substantiate this claim with 'hard' data has therefore assumed that ordinary

mature language-users analyze incoming utterances *syntactically*, and, whether or not they do it the way transformational grammar does it, they come up with the same result, *i.e.*, the 'correct' structural description for that utterance.

In view of past experimental outcomes, however, what has been to transformational grammarians an intuitively obvious fact, and to psycholinguists a tacit assumption, may not be entirely relevant. Cohen has suggested that language-users seldom, if ever, need the full structural description of any incoming linguistic waveform: a partial description of it, 'filed' under familiar cognitive categories, suffices, along with appropriate 'cross-references' through which the actual utterance can be reconstructed, if need be. The proposal is highly reminiscent of Miller's 'kernel plus syntactic footnote' recoding (1962, pp. 760-1), and bears a striking resemblance to Fodor & Garrett's 'reconstruction of deep structure' (1967, p. 290). But neither label really fits, because Cohen had something other than simplified syntactic analysis in mind: he was, in a tentative way, trying to account for "the fact that a normal hearer seems to look for sense first, and only on later reflection, if he fails to understand something, gets worried about ungrammaticalness and ambiguity" (Lyons & Wales, 1966, p. 172).

It seems eminently reasonable that linguistic

processing should involve both syntactic and semantic abilities, in greater or lesser proportion. The 'syntax-only' approach, ignoring meaning, raised far more questions than it answered, and a 'semantics-only' approach seems equally unpromising. If one assumes, as did Cohen, that hearers recode waveforms primarily, though not entirely, in terms of meaning, many anomalous experimental results can be easily explained (see Lyons & Wales, 1966, pp. 170-1). However, any attempt to justify such an assumption empirically leads only to a further dilemma: since syntactic patterns themselves presumably signal meanings, how can one ever manipulate syntax and semantics independently, let alone distinguish which is primary? Lexical variation is easily achieved through simple word-substitution in fixed syntactic frames, but syntactic structure and sentential meaning appear to be so interwoven as to preclude varying syntax while holding semantics constant.

What seems to have been overlooked for many years is that the meaning of a sentence is not merely the aggregate of the meanings of its lexical items and their semantic inter-relations; the order in which items of certain kinds occur conveys additional meaning, as do other "formal features of grammar" (Bloomfield, 1933, pp. 166ff.). And since it is the meaning conventionally associated with a pattern that has relevance for a language-user, not the pattern itself, there is really no such thing in ordinary

language-use as semantically empty syntax. Two kinds of meaning are thus distinguished: that associated with the lexical content of a sentence [semantics of content, or S(c)], and that associated with its syntactic configuration [semantics of type, or S(t)]. It is these which are the effective independent variables in any sentence-complexity study, and they are independently manipulable, whereas syntax and semantics _per se_ are not (Baker, Prideaux, & Derwing, in press).

The failure to make this distinction must surely be held to blame, along with the methodological and foundational inadequacies described above, for the conflicting interpretations of experimental results to date. In line with the traditional linguistic distinction between grammar and lexis, most workers seem not only to have assumed that syntactic and semantic processing are performatively independent, but have generally tended to ignore the latter altogether. While it is entirely legitimate for linguistic analysts to speak of pure syntax on the one hand versus semantics on the other, the dichotomy seems to have no relevance for the ordinary language-user. What he is engaged in, when he converses, is some highly complex interplay of _lexicosemantic_ and _syntactosemantic_ processing. The present study assumes such to be the case, and addresses itself to the elucidation of the latter mode of experience, for it, presumably, is what is involved in the apprehension of sentence-types.

Because of faulty initial assumptions and the consequent confounding of variables, it is difficult to reconstruct from past experimental results a coherent view of the behavioral correlates of even the most fundamental syntactic properties of sentences. Those experiments were designed mainly to estimate the relative complexities of various sentence-types, and the syntactic differences among types were tacitly assumed to be as obvious to Ss as they were to the Es. Such an assumption now appears unwarranted, for unless a language-user has learned to associate different S(t) meanings with different syntactic patterns, he will not even perceive instances of those patterns as different sentence-types, let alone respond in measurably altered fashion. Any further progress in this line of investigation would thus seem to depend on close re-examination of some rather basic premises.

The question might now justifiably be asked, is the ordinary language-user at all aware of those supposedly universal "elementary sentence types" which are so obvious to linguists (see Katz & Postal, 1964, pp. 72-3, 118-9), and which have found a place in every grammar since Plato? That the superficial reflexes of sentence-type features are perceptible is not at all in doubt. What needs to be demonstrated is whether these 'surface structures', disregarding lexical content and extrasentential factors, do indeed signal differentiated sentential meanings. Does the

language-user, in other words, associate different S(t) meaning-components with different structural patterns, regardless of the lexical burden - S(c) - an utterance might happen to carry?

<u>Sentence-Properties: a New View</u>

Most of the studies dealing with sentential complexity have employed some or all of the eight basic sentence-types in what Clifton & Odom (1966) have called 'the P, N, Q sentence-family':

1. Kernel	5. Passive Negative
2. Passive	6. Passive Question
3. Negative	7. Negative Question
4. Question	8. Passive Negative Question

It was implicitly assumed, by both linguists and psychologists, that 'passiveness', 'negativeness', and 'interrogativeness' were additive attributes that were simply either present in or absent from utterances, in various combinations, at the option of the utterer. Kernel sentences were the result of 'no optional transformations', and the syntactic (as well as the perceptual) complexity of a sentence was taken to be a function of the number of such transformationally-added properties it embodied. The <u>Aspects</u>-type grammar (Chomsky, 1965), it should be noted, remains unchanged in this fundamental respect from that of

Syntactic Structures vintage (Chomsky, 1957): passivization, interrogation, and negation are still formalized as 'yes/no' options, albeit in deep structure rather than in transformations, with non-passive, non-negative non-questions resulting from three 'no' decisions. "That is, the [simple active affirmative declaratives] are the '*faute d'autre*' sentences of English", and the current Chomskyan grammar model still incorporates the notion of "cumulative complexity" (Watt, 1970a, pp. 56-7).

It seems to have escaped general attention that, if a sentence is not a question, it cannot be just a 'non-question'; it must be something else, either a declaration or a command. If no sentential negation is being expressed, then the sentence must be affirmative. The only alternative to passive focus (see Chambers, 1970) is active focus. There are, in other words, no 'take-it-or-leave-it' options in English, as far as sentence-type features are concerned, only 'take-one-or-the-other'. Three inherent S(t) dimensions have been postulated, each dimension encompassing at least two mutually exclusive S(t) attributes; ignoring imperatives, these are:

1. *Voice* (active or passive)
2. *Mood* (declarative or interrogative)
3. *Modality* (affirmative or negative)

(Baker, Prideaux, & Derwing, in press). The eight basic sentence-types of interest, then, whatever additional

features they may ultimately incorporate, can be described as follows:

1. Active Declarative Affirmative (ADF)
2. Active Declarative Negative (ADN)
3. Active Interrogative Affirmative (AIF)
4. Active Interrogative Negative (AIN)
5. Passive Declarative Affirmative (PDF)
6. Passive Declarative Negative (PDN)
7. Passive Interrogative Affirmative (PIF)
8. Passive Interrogative Negative (PIN)

Clearly, this molecular view embodies no _a priori_ assumptions whatever about relative complexity. However, the experimental question now becomes a more complex one, for the various aspects of sentence-meaning must, like the syntactic patterns which convey them, be assumed to occur conjunctively, in at least three dimensions. The issue, then, is whether the various combinations of fundamental syntactic properties are functionally distinct to ordinary speakers of English. That is, can sentences exhibiting the same set of type-feature manifestations be consistently perceived as members of the same type-class, irrespective of lexical content, and as different from all sentences having other combinations of type-features?

An exploratory computer-controlled study [Baker, Prideaux, & Derwing, in press (henceforth 'BP&D')] has shown

that naive speakers of English can indeed learn to recognize an instance of any one of the eight basic sentence-patterns, despite extensive variation in lexical content. Distinctions, furthermore, were not made entirely on the basis of the overt syntactic reflexes of the co-occurring type-features; the data strongly indicated that "it was the semantic significance [S(t)] of different types, rather than pattern differences as such, to which a subject responded".

The 32 BP&D Ss, however, were a relatively small, heterogeneous sample, undifferentiated with respect to sex or linguistic sophistication, and they varied widely in the number of trials they required to accomplish the learning task. Furthermore, the sentences to be distinguished were presented in the visual mode, which linguistics traditionally regards as ancillary to the oral-aural communication channel. The feasibility of the BP&D experimental paradigm was, however, amply demonstrated, and the present study was an attempt to replicate their results in the aural perception mode, with a larger and more homogeneous sample of Ss. These were systematically differentiated as to sex; no real performance-differences were expected, in the perception of syntactic cues, but only through incorporating sex as an independent variable could an empirical estimate be obtained of the relative importance of this factor (see Hays, 1963, p. 450). Other modifications of the BP&D procedure, such as equalizing the number of trials for all Ss, are outlined more fully below.

The Experiment

Conjunctive Concept-Formation

Any behavioral test of the experimental question outlined above must conform to certain requirements; these, in turn, are dictated by the nature of the question itself and of its concomitant assumptions. All sentence-types must, first of all, be equally available to observation; yet, at the same time, their assumed three-dimensional syntactic composition must be accessible to analysis by Ss, and reflected in some correlated dependent variable. A conjunctive concept-formation (CCF) task seemed the most appropriate method of investigation: it was this experimental paradigm which provided data for the BP&D study, and it did so for the present one, as well.

In both experiments, Ss were presented a number of sentences in which the three syntactic dimensions, as the main independent variables, had been manipulated orthogonally; the sentences also differed systematically with respect to certain 'distractor' variables. The task was for Ss to discover, despite irrelevant variations in sentence form and meaning, that 'secret' subset of the sentences which embodied one or another conjunctive combination of the type-features. The dependent variables of principal interest were the number of trials required for each type-feature to be adverted to, the order in which they

were 'discovered', and the distribution of the classification errors. These served to suggest inferences about the relative perceivability of each syntactosemantic type-feature.

The methodology of concept-formation ('concept-learning') is well-established in the investigative repertoire of the psychology of learning (Deese & Hulse, 1967, Ch. 12). In its 'response method' application, it is regarded as an extension of the standard paired-associates technique: where the latter promotes the (rote) acquisition of associations on a one-to-one basis, the former seeks to establish a single many-to-one correspondence between an abstract (but often describable) target-class and a number of positive instances of that class's membership. The response method itself involves the serial presentation of stimulus-objects in random order, with the learner attempting in each trial either to name the abstract category of which the given stimulus is a putative member, or, as in the present study and its antecedent, to judge whether it is or is not a member of the target-class. In either case, the E immediately reinforces the S's judgment positively or negatively, depending on

1. whether the given stimulus was a positive instance of the target-class, and
2. whether the S's judgment was 'yes' or 'no'.

Negative reinforcement is generally put to use by Ss,

through some individualized learning strategy, to eliminate wrong hypotheses about the relevant attribute(s) of the target-class; positive reinforcement serves, of course, to confirm correct hypotheses. Eventually, Ss come to respond correctly to all stimuli, and can be said, after a sufficient number of consecutive correct judgments, to have 'achieved criterion', i.e., formed the abstract concept. If they can then also describe the target-class verbally, they are said to have learned the concept as well (Hunt, 1962).

Target concepts are defined in terms of those stimulus properties which are under investigation. Of course, Ss are not informed in advance of the relevant attributes which delimit their target; it is their task to discover them, by trial and error. Concept-definition may be disjunctive, involving only one of the attributes in some dimension of interest, or, as in the present studies, conjunctive, where one attribute in each of two or more descriptive dimensions simultaneously define the target-class.

Applying the CCF Method

Besides seeming admirably suited to the phenomenon under investigation here, the CCF method had the further advantage of fulfilling the need for a minimally memory-dependent assessment technique, argued for above (see 'Methodological Problems'). In the present experiment, as

in the BP&D study, the rate of presentation of stimulus-sentences was controlled by the Ss; this obviated the need for memorization, and provided the opportunity for as full a consideration of each sentence as Ss might require (see Fillenbaum, 1970, p. 236). In order, furthermore, to inhibit stimulus-specific memory - and thus make lexicosemantic considerations irrelevant to the task - both studies implemented the following measures:

1. Only 16 lexical-content sets were employed, each developed into all eight basic sentence-types; the 128 resultant sentences were thus unique, but would not seem so to Ss, especially with aural presentation. As Miller & McKean have pointed out, the repetition of subject matter requires Ss to attend more to grammatical structure (1964, p. 299).
2. The 16 lexical-content sets were differentiated with respect to the constitution of their agent ('actor') and patient ('acted-upon') Noun-Phrases, and of their auxiliary-verb phrases as well, in a manner uncorrelated with the variations in the main type-features. Any attempt by Ss to use tense, number, sentence-length, etc. as classificatory criteria thus went unrewarded; concepts could only be formed by ignoring non-S(t) attributes.

The use of complex linguistic signals in concept-formation constitutes something of a departure from familiar procedure, particularly in the case of aural presentation.

It opens the door to certain methodological problems which must be accounted for. First, highly structured stimulus-objects such as sentences are virtually unlimited, as far as perceivable properties are concerned; such a large number of necessarily irrelevant attributes might tend to inhibit concept-attainment and make data-interpretation rather difficult. Furthermore, the high proportion of non-target categories, in relation to the single target-class for every S, makes for an unusually large ratio (8:1) of negative to positive instances in the stimuli; this situation could make negative instances rather less informative to Ss, thus complicating their task.

Such doubts may validly be raised, in the light of traditional CF experimentation. But the BP&D results showed these fears to be largely unfounded, and enabled the present study to be undertaken without trepidation. Judging from their response-patterns and later verbal reports, the Ss in both experiments tended, by and large, to employ the 'wholistic conservative focusing' strategy, which has been found to be most common, and also optimal. This is the approach where Ss test hypotheses about one stimulus-attribute at a time, eventually narrowing the choice down to some unique combination of attributes (Deese & Hulse, 1967, pp. 423-4). This approach was in fact suggested to the Ss of both studies in their pre-test instructions, and their subsequent performances did not seem to be at variance with

traditional expectations.

Although a number of Ss in the BP&D study did fail to attain the required concept within a reasonable time, and those who did varied widely in their 'trials-to-criterion' (TTC) scores, this gross performance measure turned out to be of but marginal interest. Ss' yes-response distributions over trials and with respect to stimulus-type were far more suggestive of ongoing psycholinguistic processes, namely, of the primacy of S(t) and of the subjective dominance-hierarchy among the effective stimulus-properties (Deese & Hulse, 1967, pp. 433-4). The present experiment therefore ignored TTC, presenting all sentences once each to all Ss. Since 94% of the undergraduate Ss in a visual-mode pilot study had achieved criterion within 128 trials, few, if any, in the aural version were expected to experience difficulty in doing so. The main data then became the patterns of response, rather than the number of trials required to reach some necessarily arbitrary criterion.

As to the relative rarity of positive instances, if this tended to prolong concept-attainment, it was of little consequence, for TTC was no longer a dependent variable of interest. Besides, it has been demonstrated that "people can learn to deal with negative instances about as efficiently as they can with positive instances" (Deese & Hulse, 1967, p. 425). In any case, there is no other way of incorporating a 2^3 factorial design into the CCF paradigm;

if a greater tendency towards classification errors is the result, this is no hindrance, for it is their distribution with respect to stimulus-type that is crucial, not their absolute number.

CHAPTER TWO

METHOD

<u>Stimulus-Materials</u>

The sentences presented to <u>S</u>s for categorization were a set of 128 distinct strings, formed by combining 16 different lexical content-sets with the syntactic frames for the eight basic sentence-types (see p. 26, above). Each content-set consisted of an agent Noun-Phrase, a transitive verb, and an object Noun-Phrase; the following are examples:

(01) The boy . . . help . . . the girl.

(06) Many children . . . throw . . . snowballs.

(12) Paul . . . entertain . . . our guests.

The Noun-Phrases (NPs) were of seven structural types, and each content-set incorporated one or another pairing of these types; a full tabulation of the NPs employed will be found in Appendix A.

The sentences were cast in four different traditional 'tenses': present-progressive, past-progressive, perfect, and future. This was effected by incorporating the aspectual elements <u>be</u> (for progressive) or <u>have</u> (for

perfective), or the modal will as auxiliary verbs, with the appropriate adjustments in the verb-phrases. Thus, every sentence was derived from one of the following patterns:

1.	[NP^1]	[is/are	Verb + ing]	[NP^2]
2.	[NP^1]	[was/were	Verb + ing]	[NP^2]
3.	[NP^1]	[has/have	Verb + en*]	[NP^2]
4.	[NP^1]	[will	Verb]	[NP^2]

However, each of the 128 possible 'content-set by syntactic type' combinations appeared in only one tense, for, within the eight syntactic realizations of any content-set, only those pairs of sentences differing in all three type-dimensions were in the same tense, and, at the same time, only four of the 16 possible sentences of each syntactic type were in the same tense. Tense was thus distributed among the 128 sentences according to an arrangement of four 4 X 4 Latin squares and their reflections; this distribution is displayed in Appendix B. Appendix C contains a complete listing of the 128 sentences, in basic presentation order (described below).

The variations, systematic cr otherwise, in NP structure, tense, and lexical content constituted irrelevant 'distractor' characteristics of the sentences, for the CCF task. They were presumably available to Ss as

---

* The en morpheme represents the past participle suffix appropriate to whatever verb it occurs with, e.g., '-en' itself for eat, but '-ed' for fix.

classificatory criteria, among any others they might perceive or imagine, but Ss had to learn, by trial and error, to ignore them, in order to focus on the relevant Voice, Mood, and Modality properties.

Subjects and Procedure

Sixty-four undergraduate students at the University of Alberta served as Ss, 4 males and 4 females in each of the eight target-groups. They volunteered their help in response to a general appeal through posters; a monetary reward was offered for the best three performances in each block of 32 Ss. One volunteer had to be rejected in mid-test, when it became evident that the instructions had been completely misunderstood. Ss were seated at a small table, holding a remote-control button by means of which tape-recorder playback could be (re)started. Instructions and stimulus-sentences issued from an extension speaker on the table in front of the S. The E sat behind and to the left of the S.

The pre-experimental instructions to Ss (see Appendix D for the full text) had been recorded by the E and his wife; they were presented to Ss in two parts, as follows:

1: an explication of several properties in terms of which otherwise non-identical sentences might be categorized as of the same 'type': the properties suggested and exemplified included tense and content, as well as the

relevant ones for the experiment, both conjunctively and disjunctively;

2: details of the experimental procedure: to each sentence-presentation, Ss were to respond 'yes' or 'no', reflecting their judgment on that sentence's membership in the (unspecified) target-class; each such response would be immediately reinforced by the E as 'correct' or 'wrong', thus enabling Ss to learn to distinguish relevant from irrelevant properties; during the initial trials, when nothing at all could be known about the target sentence-type, Ss were to respond 'no' consistently, until they were told that this response was 'wrong', indicating that they had just heard (and inevitably missed) a target-type sentence; from then on, Ss were to respond as they saw fit.

The tape-playback was halted by the E after each section of the instructions, so that Ss could ask for further clarification, if they wished to; Ss then restarted the tape player themselves, whenever they were ready to proceed.

The presentation of the 128 stimulus-sentences (recorded by the E) was self-paced by Ss. After each sentence, the E halted the tape player, so that the S could respond at his convenience, then receive immediate reinforcement; when the S was ready for the next sentence, he restarted the tape himself. A silent gap of about three seconds separated the sentences on the tape. All S-responses ('yes' or 'no') and all E-reinforcements

('correct' or 'wrcng') were recorded on scoring-sheets for later processing.

For every S, the first occurrence of a target-type sentence was in his fourth trial. In order to effect this, the 128 sentences were arranged such that the fourth sentence in each of the eight successive blocks of 16 was of a different target-type, and was preceded by three different non-target-type sentences. Every S thus entered the overall sentence-sequence at one of eight points, depending on his assigned target condition. All Ss were accorded equivalent treatment in their first four trials, in that they heard:

(1st) one of the 16 Mood+Modality opposites of their target-type;

(2nd) one of the 16 Voice-opposites cf their target-type;

(3rd) one of the 16 Voice+Mood+Modality opposites of target-type;

(4th) one of the 16 of their target-type.

Since all Ss had been instructed to respond 'no' until negatively reinforced, these first four trials were ignored, in the subsequent data-analyses.

The 128 sentences were presented once each, in cyclic fashion (see Appendix C). As a result of the above arrangement, each target-condition was embedded in a distinct presentaticn-crder. Eight such presentation-orders had to be re-reccrded from the master tape, one for each

target-condition.

All sentences of predetermined type were, of course, randomly selected from the 16 available in each type-category. The remaining sentences in the basic sequence were randomly selected with respect to both type and content, with the following restrictions:

1. No content-set was permitted to occur twice in succession;
2. No syntactic type was permitted to occur twice in succession; at least two other syntactic types had to intervene;
3. Each of the 16 successive blocks of eight had to contain one of every syntactic type, in a different order;
4. Each of the eight successive blocks of 16 had to contain one of every lexical content, in a different order.

The purpose of these restrictions on random presentation was to ensure an equitable distribution of positive and negative instances throughout the experimental session for every S.

Ss were randomly assigned to target-groups. Experimental sessions were timed, in order to provide an ancillary (hopefully covariant) dependent measure of overall response rate.

TABLE 1

DATA STRUCTURES FOR THE EXPERIMENT

			Stimulus - Features							
			Active				Passive			
			Declar.		Interr.		Declar.		Interr.	
Sex	Tgts.	Ss	Aff.	Neg.	Aff.	Neg.	Aff.	Neg.	Aff.	Neg.
		1								
M	1 ADF	2								
		3								
		4								
.	.	.								
.	.	.								
.	.	.								
		29								
M	8 PIN	30								
		31								
		32								
		33								
F	1 ADF	34								
		35								
		36								
.	.	.								
.	.	.								
.	.	.								
		61								
F	8 PIN	62								
		63								
		64								

Re-organization of Data within Ss:

Corr. Resp.	Voice Err's	Mood Err's	Modal. Err's	Voice +Mood Err's	Voice +Mod. Err's	Mood +Mod. Err's	Voice +Mood +Mod. Err's

Design

Ss were grouped in terms of two crossed factors - sex (two levels) and target-type (eight levels) - with four Ss

within each subgroup. Sixteen sentences of each target-type were presented to all Ss in essentially random order for 'yes/no' judgment. Since the eight sentence-types were defined in terms of three crossed dichotomous factors, Voice, Mood, and Modality, the basic experimental design was a five-factor factorial, with repeated measures on the last three (treatment) factors, and Ss within Sex X Target subgroups as replications (see Table 1).

It has been noted that each target-condition was embedded in a distinct presentation-order; this forces the assumption of a nonsignificant order effect. This assumption is not unwarranted, since every S heard every sentence in what must have been to him a totally unpredictable pattern, with respect to the occurrence of syntactic types, tenses, contents, and every other conceivable property. Furthermore, all Ss were reinforced at the same rate, *i.e.*, one positive instance of the target-class in every block of eight trials.

The dependent variable of primary interest was the number of yes-responses made to each stimulus-type. Since 'yes' was erroneous after all but target sentences for each S, the 'yes-counts' in all non-target categories were interpreted as indices of the confusibility of the various non-target sentence-types with Ss' constantly changing notions of their target-type. Any significant deviation of these 'yes-error' counts, across stimulus-types, from a

uniform chance distribution would tend to support the notion that the eight basic sentence-types are not equally confusible with each other.

However, since what constituted 'target' or 'non-target' varied from one target-group to the next, so did the implications of the yes-error distributions. Thus, if inter-type (dis)similarities are correlated with differences in Voice, Mood, or Modality, the factorial arrangement outlined above would tend to obscure this fact. The structure shown in the main body of Table 1 therefore provided only a starting-point; analyses of CCF-strategy and dominance-hierarchy necessitated the re-ordering of cells within rows into the arrangement shown at the foot of the table. In this array, all entries in the first column represent counts of yes-responses made (correctly) to target-type sentences, whichever they were for any S. Entries in the second column represent counts of erroneous yes-responses made to sentences which differed from that S's target-type in the Voice dimension only. The remaining columns contain counts of erroneous yes-responses made to sentences which deviated from target-type in Mood, Modality, and the various combinations of the three feature-dimensions. The rightmost column subsumed yes-responses to sentences of completely opposite type.

'No'-responses perforce went unanalyzed. In view of the explicit instructions, it was assumed that Ss said 'yes'

only to what they thought were target-type sentences; but no assumption whatever could be entertained with respect to 'no' responses. These, if Ss performed as instructed, were given after sentences which were thought not to be target, or whose status was not apparent.

CHAPTER THREE

RESULTS

Effects of Subject Characteristics

Table 2 displays the correlations among three *S*-attributes (sex, age, and year of university) and three behavioral variables in the experiment: total time (in min.) taken to process the 128 sentences, total number of yes-responses made, and total number of errors in sentence-classification. The means and standard deviations (SDs) of these variables are also shown. With 63 degrees of freedom (*df*), correlations as low as 0.25 are regarded as statistically greater than zero; but such are sufficiently less than unity to warrant cautious interpretation of their practical significance. The trends that did emerge, however, were not altogether unexpected.

The distinct lack of correlation between sex and any of the dependent variables was not what one would expect, in the light of traditional views on language skills. But there is no real reason to believe that males should perceive syntactic cues differently from females, so this

outcome was therefore not surprising. What it suggests is that, in the present type of experimentation, at least, sex need not be considered a relevant source of variation. This matter is taken up again in Chapter Four.

TABLE 2

ORGANISMIC AND DEPENDENT VARIABLES

	Variables					
	Sex	Age	Year	Time	Yeses	Errors
Means:		20.52	2.25	23.53	17.73	12.80
Std.Deviations:		2.54	0.92	4.5	4.39	8.62
Correlations:						
Sex	1.0					
Age	0.10	1.0				
Year	0.14	0.17	1.0			
Time	0.04	0.56	0.01	1.0		
Yeses	-0.16	0.04	-0.22	0.12	1.0	
Errors	0.09	0.33	0.0	0.37	0.35	1.0

Another correlation, that between age and year of university, appeared conspicuous by its non-significance, for one would think that older students would tend consistently to be more advanced. But the undergraduate sample drawn for this study seems to have incorporated an unusual degree of variability in this regard, as the SDs in Table 2 will attest. The 64 Ss appeared on the whole to have been 'normal' undergraduates (mean age (in spring term) = 20.5, mean 'year' = 2.25), but only 43 of them were in

fact within one year of the 'normal' age/years gradient; Ss ranged in age from 16 to 32, and both these extremes were freshmen. The removal of the latter S from the analysis yielded an age:year correlation more in line with expectations ($r = 0.325$; $p < 0.01$). The 'undergraduate' criterion does not seem to make for as homogeneous a sample nowadays as it once might have.

Age seems to have been the only organismic variable to correlate with any dependent variable: older Ss took more time to complete the task ($r = 0.56$; $p < 0.01$), and, like most slow responders in timed tasks, tended to make more errors ($r(time:errs) = 0.37$, $r(age:errs) = 0.33$; $p < 0.01$). These correlations were somewhat diminished by the omission of the oldest S, but remained statistically significant, nevertheless. Since the sample included a number of older Ss, a trend towards the deliberateness characteristic of the more mature was not surprising. Their proneness to error, however, is another matter, and could only be speculated upon, at this point.

An analysis of variance (complete factorial with replications) on the time measures revealed no significant differences whatever among target-groups, with respect to this variable. The presence of a highly deviant time-score, 46 min., (contributed by the oldest S), made for a rather large estimate of error-variance. But it appears unlikely that any of the effects would have turned out significant,

in any case; one can only conclude that this measure was too gross to reflect performance-differences as a function of target-type.

The one remaining 'significant' correlation was that between yes-counts and errors ($r = 0.35$; $p < 0.01$). This was in part due to the non-independence of these two counts, for a yes-response could be either correct or wrong. But in addition to that, because the number of target-type sentences for every S̲ was 16 of the 128, the probability of a 'yes' response being wrong was eight times greater than that of a 'no' being erroneous, on the basis of chance alone. The more a S̲ tended to respond 'yes', therefore, the greater his error-count was likely to be.

Distribution of Yes-Responses

The most gratifying result of this experiment was the degree to which its yes-response distribution coincided with that obtained by BP&D in their visual CCF study of subjective syntactic features. Table 3 shows the cell means of both experiments, averaged over male and female S̲s, and expressed as a percentage of trials, for comparison; means within target-groups are arranged according to stimulus-type responded to. The corresponding cell values of the two experiments were highly correlated ($r = 0.896$), providing rather convincing evidence that the present experiment had replicated the BP&D results.

TABLE 3

MEAN PERCENT YES-RESPONSES TO STIMULUS-TYPES
AS A FUNCTION OF TARGET-TYPE

		Stimulus-Sentence Types							
Target Type	Expmt.	ADF	ADN	AIF	AIN	PDF	PDN	PIF	PIN
ADF	BP&D	7.38	1.78	0.58	0.82	5.20	1.29	0.85	0.29
	REID	6.75	0.60	0.71	0.40	3.63	0.91	0.60	0.20
ADN	BP&D	0.51	7.18	0.38	1.01	0.44	4.09	0.58	0.67
	REID	0.40	8.27	0.20	0.10	0.60	3.02	0.10	0.50
AIF	BP&D	0.59	0.88	5.56	3.16	0.44	0.29	3.60	1.65
	REID	0.30	0.00	9.68	1.41	0.50	0.30	2.22	0.60
AIN	BP&D	0.53	0.92	1.98	7.84	0.35	0.00	0.14	3.11
	REID	0.20	0.81	0.81	9.17	0.20	0.50	0.91	3.43
PDF	BP&D	3.37	0.58	0.41	0.21	7.92	0.62	1.11	0.00
	REID	2.12	0.10	0.50	0.10	9.27	1.21	0.91	0.30
PDN	BP&D	1.02	6.05	0.89	2.60	1.71	8.26	1.17	2.20
	REID	0.71	1.81	0.00	0.50	1.61	7.36	0.20	0.50
PIF	BP&D	2.05	0.48	3.43	2.60	1.57	1.24	7.97	2.17
	REID	0.40	0.20	2.52	1.81	1.21	0.20	5.44	3.33
PIN	BP&D	1.26	0.95	1.49	4.37	0.74	1.06	1.24	5.40
	REID	0.00	0.30	0.30	4.03	0.00	0.50	0.50	8.37

The majority of the Ss' yes-responses in the present study were to sentences of target type, as inspection of Table 3's main diagonal will reveal. This simply means that the conjunction of syntactic stimulus-features which came to elicit the most consistent acquiescence was that whose

recognition was being positively reinforced. It is not surprising, therefore, that the most significant effect in the preliminary analysis of variance was the Targets X Stimuli interaction ($\underline{F} = 23.53$; $p < 0.001$). The means for yes-response among stimulus-categories were the only other significant source of variance ($\underline{F} = 2.70$; $p < 0.01$), suggesting that some sentence-types were more often thought to be target, right or wrong, than others. But this factor was probably of marginal practical significance, for the proportion of variance accounted for was less than 1% (see Hays, 1963, pp. 409-10). Besides, the pronounced Targets X Stimuli interaction denies the existence of an autonomous stimulus-features effect: what differences there were, in yes-response to various stimulus-types, were dependent on target-assignment, and the analysis of errors below makes this readily apparent.

Table 4 contains the cell means for yes-response counts, distributed in terms of sex, target-type, and error-type. This table was derived from the basic data-arrangement in the manner described in Chapter Two, under 'Design': the leftmost column subsumes all correct yes-responses (i.e., those made to target-type stimuli); Ss' remaining yes-counts are distributed through the other columns as a function of stimulus deviation from their respective target-types. The appropriate analysis of variance is summarized in Table 5, and the complete table,

TABLE 4

YES-RESPONSE MEANS
AS A FUNCTION OF ERROR-TYPE

			Types of Error						
Tgt	Grp.	Correct Resp.	Voice	Mood	Mod.	Voice Mood	Voice Mod.	Mood+ Mod.	Voice+ Mood+ Modal.
	Male	6.0	4.5	0.75	1.0	0.25	1.75	0.75	0.5
ADF	Fem.	10.75	4.5	1.0	0.5	1.25	0.5	0.25	0.0
	Mean	8.375	4.5	0.875	0.75	0.75	1.125	0.5	0.25
	Male	9.25	3.5	0.0	0.25	0.5	0.5	0.0	0.0
ADN	Fem.	11.25	4.0	0.25	0.75	0.75	1.0	0.5	0.25
	Mean	10.25	3.75	0.125	0.5	0.625	0.75	0.25	0.125
	Male	12.5	3.25	0.5	1.0	1.25	0.75	0.0	0.5
AIF	Fem.	11.5	2.25	0.25	2.5	0.0	0.75	0.0	0.25
	Mean	12.0	2.75	0.375	1.75	0.625	0.75	0.0	0.375
	Male	13.0	4.0	1.75	0.75	1.0	0.25	0.5	0.5
AIN	Fem.	9.75	4.5	0.25	1.25	0.25	2.0	0.0	0.0
	Mean	11.375	4.25	1.0	1.0	0.625	1.125	0.25	0.25
	Male	10.0	1.25	1.5	2.25	1.25	0.0	0.75	0.25
PDF	Fem.	13.0	4.0	0.75	0.75	0.0	0.25	0.0	0.0
	Mean	11.5	2.625	1.125	1.5	0.625	0.125	0.375	0.125
	Male	6.5	2.75	0.5	4.0	0.5	1.5	0.5	0.0
PDN	Fem.	11.75	1.75	0.75	0.0	0.75	0.25	0.0	0.0
	Mean	9.125	2.25	0.625	2.0	0.625	0.875	0.25	0.0
	Male	6.75	2.5	1.25	3.75	0.5	1.75	0.25	0.25
PIF	Fem.	6.75	3.75	1.75	4.5	0.5	2.75	0.25	0.25
	Mean	6.75	3.125	1.5	4.125	0.5	2.25	0.25	0.25
	Male	9.75	2.75	1.0	0.75	0.25	0.25	0.0	0.0
PIN	Fem.	11.0	7.25	0.25	0.5	0.5	0.5	0.0	0.0
	Mean	10.375	5.0	0.625	0.625	0.375	0.375	0.0	0.0
Means:									
	Male	9.129	3.063	0.906	1.719	0.688	0.844	0.344	0.25
	Fem.	10.719	4.0	0.656	1.344	0.5	1.0	0.125	0.094
	Grand	9.969	3.531	0.781	1.531	0.594	0.922	0.234	0.172

showing the subcomponents of the main effects and interactions, is presented in Appendix E. The absence of significant variation attributable to sex or target-group assignment indicates that yes-responses were in no way dependent on these factors; the assumed conflatability of between-S factors thus appears to have been justified (see 'Design', Chapter Two).

TABLE 5

ANALYSIS OF VARIANCE:
YES-RESPONSES AS A FUNCTION OF SEX, TARGET, AND ERROR-TYPE

Source	SSq	df	MSq	F
Sex (S)	3.955	1	3.955	1.613
Target-types (T)	12.764	7	1.823	0.744
S X T	19.873	7	2.839	1.158
Ss(S X T)	117.717	48	2.452	
Error-types (E)	4913.543	7	701.935	156.687**
S X E	51.465	7	7.352	1.641
T X E	335.750	49	6.852	1.530*
S X T X E	221.779	49	4.526	1.010
Ss(S X T X E)	1505.228	336	4.48	

* $p \leq 0.05$ ** $p \leq 0.01$

The only effect of unquestionable significance would seem to be the error-types component (F = 156.69; $p <$ 0.001). The Sex X Target-mood X Errors interaction - part

of the otherwise inconsequential Sex X Targets X Errors component (see Appendix E) - can be ignored altogether, for reasons just cited above, and the same applies to the Sex X Target-voice X Target-mood interaction. Even the barely significant Targets X Errors interaction (df = 49, 336; $p < 0.05$), and its 0.01-level subcomponents must not be overestimated, since they account for less than 2% of the total variance. A graph of this interaction (Figure 1) makes it readily apparent that only one target-group - the PIF - deviates from the general error-response pattern: despite the fact that the PIF mean is not significantly different from the other target-group means, the group yes-response profile deviates from the rest enough to create a marginal interaction effect. This group thus seems to be responsible for the noticeable Targets X Errors component, and for some other ancmalies which are discussed below.

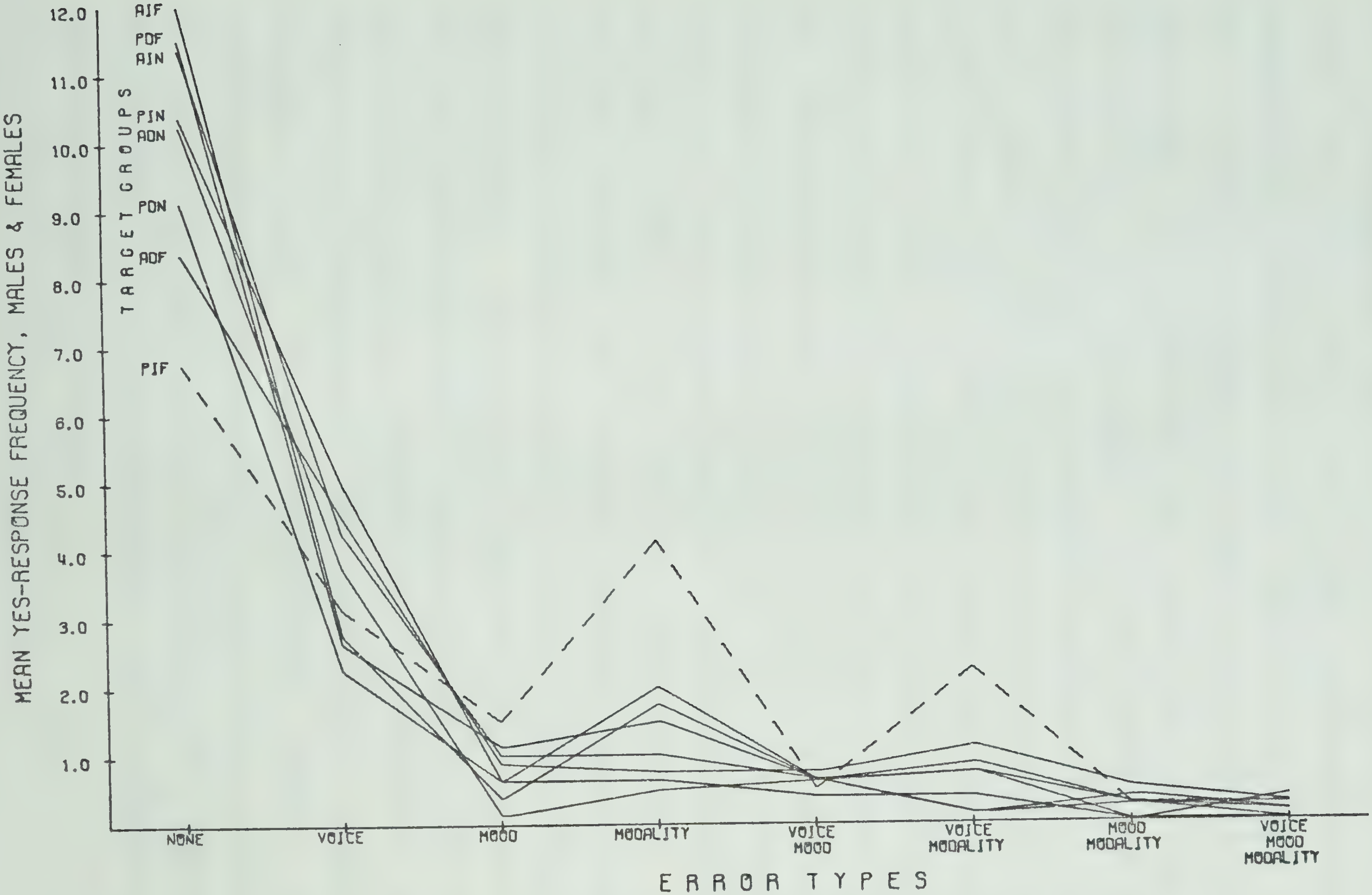

FIGURE 1. TARGETS X ERRORS INTERACTION

An *a posteriori* Tukey test among all error-type means (Table 6; see Winer, 1962, p.87) revealed further correspondence between the present results and those of BP&D. The means themselves fell into the same order in both experiments, and exhibited very nearly the same inter-relationships. The preponderance of yes-responses was, as has already been shown, to target-type sentences; it is thus to be expected that the mean for correct responses in the present study should differ significantly from all other means ($p < 0.01$). Voice-errors, while significantly less frequent than correct yes-responses, were more common than any other type of error ($p < 0.01$). As in the BP&D study, the difficult active/passive distinction was the greatest source of false positive responses. Where the two studies diverged slightly was with respect to Modality-errors. In the BP&D results, all types of yes-error except those in Voice were of like low frequency, differing insignificantly from each other. In the present experiment, however, errors in this affirmative/negative dimension occurred in somewhat greater proportion: their incidence was still markedly less than that of Voice-errors, but greater than the least-occurring types, namely Mood+Modality errors and Voice+Mood+Modality errors ($p < 0.01$). With respect to each other, errors in Mood only (declarative/interrogative), two dimensions, or all three dimensions were statistically indistinguishable in frequency.

TABLE 6

TUKEY'S 'HONESTLY SIGNIFICANT DIFFERENCE' TEST
AMONG ALL MEANS FOR ERROR-TYPE

Means for Error-type	Differences Among Error-type Means						
	Vce+Md. +Modal. 0.172	Mood+ Modal. 0.234	Voice+ Mood 0.594	Mood 0.781	Voice+ Modal. 0.922	Modal. 1.531	Voice 3.531
Correct (9.969)	** 9.797	** 9.735	** 9.375	** 9.188	** 9.047	** 8.438	** 6.438
Voice (3.531)	** 3.359	** 3.297	** 2.938	** 2.750	** 2.609	** 2.000	
Modality (1.531)	** 1.359	** 1.297	0.938	0.750	0.609		
Vce.+Mod. (0.922)	0.75	0.688	0.328	0.141			
Mood (0.781)	0.609	0.547	0.188				
Vce+Mood (0.594)	0.422	0.359					
Mood+Mod. (0.234)	0.063						

** $p \leq 0.01$

<u>Concept-Acquisition Patterns</u>

In a CCF task, the occurrences of the various types of yes-error are distributed over trials, and usually in non-random fashion, such that some types of error tend to occur only in the earlier part of the experimental session, if at all, while others persist into later trials. The last trial in which a given type of error occurred may be taken as evidence that the <u>S</u> has learned to make a distinction, and (partially) formed the target-concept. By examining the

chronology of these 'last-error-trials' (LETs), one may hope to gain some insight into the concept-acquisition strategies employed, and thereby arrive at estimates of the relative distinguishability of the syntactic features.

In the present experiment, a much greater number of Ss exhibited 'aberrant' CCF performance than had been anticipated. On the one hand, 31% of them (ten males and ten females) failed to achieve the arbitrary but useful criterion of 24 consecutive correct responses employed by BP&D, and thus cannot be said with any certainty to have formed a three-way syntacto-semantic concept by the end of the experimental session. Of the twenty, eleven had not yet learned to distinguish target-sentences from their Voice-opposites by trial 128; three were consistently making Modality-errors, and six were still responding in essentially random fashion. This 'failure' rate was rather less than that of the BP&D senior high-school Ss, 47% of whom required more than 128 trials to complete the task, but considerably more than the 6% rate among the undergraduate Ss in the visual-mode pilot study. On the other hand, 58% of the Reid Ss (19 males, 18 females) made no yes-error in one or more of the seven possible error-categories (see Table 4). Of these, twelve (19%) made no yes-error in either of two syntactic dimensions, and three (5%) made no yes-error at all. Of these three, two made no error of any kind except the inevitable 'No-wrong' in Trial 4; they had

apparently identified their target-type correctly from only one instance of it. Undergraduates in the pilot study who responded similarly amounted to 44%, 16%, and 6% of the sample, respectively, with both Ss in that 6% minority making only the one No-error.

With such a preponderance of extreme LET values entering the analysis, the resultant LET-means could be expected to embody a high degree of variance. The same was probably true of the BP&D LET-means, since nearly half of their Ss had TTC scores in excess of 128. It was not surprising, therefore, to find a barely-significant correlation of 0.48 between the two sets of LET-means, hardly enough to warrant inferring any replication effect. But the point of the analysis was to establish the dominance-hierarchy among the three syntactic dimensions. In this, the LET-scores and means themselves are rather less relevant than the rank-ordering which they indicate, of distinctions acquired over trials. Table 7 compares the orders in which Voice, Mood, and Modality differences are assumed to have been learned, by the Reid and BP&D Ss in general, and by the Reid 'pseudo-TTC' Ss (see below).

TABLE 7

MEANS AND RANKS* OF 'LAST-ERROR-TRIALS'

Tgt	Subgroup	Type of Error: Voice	Mood	Modality
ADF	Reid Ss	71^{3}	30^{2}	27^{1}
	BP&D Ss	68^{3}	22^{1}	31^{2}
	'TTC' Ss (4) **	46^{3}	32^{2}	30^{1}
ADN	Reid Ss	50^{3}	13^{1}	19^{2}
	BP&D Ss	63^{3}	34^{2}	17^{1}
	'TTC' Ss (7)	42^{3}	7^{1}	19^{2}
AIF	Reid Ss	34^{2}	12^{1}	42^{3}
	BP&D Ss	71^{3}	22^{1}	56^{2}
	'TTC' Ss (7)	37^{3}	9^{1}	36^{2}
AIN	Reid Ss	59^{3}	8^{1}	15^{2}
	BP&D Ss	48^{3}	38^{2}	32^{1}
	'TTC' Ss (7)	54^{3}	12^{2}	11^{1}
PDF	Reid Ss	40^{3}	31^{2}	28^{1}
	BP&D Ss	60^{3}	24^{2}	20^{1}
	'TTC' Ss (6)	32^{1}	35^{3}	34^{2}
PDN	Reid Ss	38^{2}	40^{3}	36^{1}
	BP&D Ss	78^{3}	55^{2}	30^{1}
	'TTC' Ss (5)	29^{3}	23^{2}	16^{1}
PIF	Reid Ss	53^{2}	37^{1}	73^{3}
	BP&D Ss	43^{3}	26^{1}	35^{2}
	'TTC' Ss (3)	49^{3}	5^{1}	40^{2}
PIN	Reid Ss	47^{3}	19^{2}	6^{1}
	BP&D Ss	72^{3}	40^{2}	38^{1}
	'TTC' Ss (5)	34^{3}	1^{1}	3^{2}
Means	Reid Ss	49^{3}	24^{1}	31^{2}
	BP&D Ss	63^{3}	33^{2}	32^{1}
	'TTC' Ss (44)	40^{3}	15^{1}	23^{2}

* Mean %'s of trials; ranks indicated by superscripts.

** Reid Ss who achieved BP&D finish criterion, with (n) of each group.

The BP&D Ss were quite consistent in making the active/passive distinction last: the mean LET for Voice in that study had the rank of 3 in all target-groups. Furthermore, where the target-type was of 'marked' Mood (AIF or PIF) or 'marked' Modality (ADN or PDN), they resolved those differences first. The distribution of mean LET-ranks with respect to error-type was significantly non-random, in the BP&D experiment ($\chi^2 = 25.5$; $p < 0.01$). Of the Ss in the present study, the females exhibited a greater consistency of approach. Like the males, they did not always advert to the Voice dichotomy last, but, where the target was interrogative (AIF, PIF, AIN, or PIN), they resolved the Mood distinction first, and for declarative negative targets (ADN or PDN), they disposed of the Modality problem first. In the last two patterns, it will be noted, the females emulated the BP&D Ss. Thus, while the overall distribution of mean LET-ranks in the present experiment did not deviate significantly from chance, that for the females alone did, beyond the 0.02 level ($\chi^2 = 11.998$).

Since twenty Ss in the present study were still committing classification errors of one kind or another at the end of the experimental session, they could not really be said to have experienced three last-error trials, as did all the rest. Their inclusion in the LET analysis was therefore regarded as a contaminating influence. In order to make a more legitimate comparison between the present

results and those of BP&D, these non-finishers were excluded from consideration, and the performances of the remaining Ss were re-examined, disregarding responses made after 24 consecutive correct ones. These pseudo-TTC Ss were then found to have responded rather more like the BP&D Ss, with all target-groups except PDF solving the Voice distinction last. The distributions of mean LET-ranks were significantly non-random for both males and females, and that for the pseudo-TTC Ss on the whole was significant beyond the 0.01 level (χ^2 = 17.25). This distribution, combined with that for the BP&D study, was systematic well beyond the 0.01 level (χ^2 = 39.75; df = 4); this strongly suggests that, if Ss are induced to form a complete three-way concept, they will agree highly on the relative perceivability of syntactic features, regardless of the perceptual mode.

Aural vs. Visual Presentation

It seems fairly clear that the once-only aural presentation of sentences posed a more difficult task for the Ss of the present experiment. This conclusion is based on comparisons with the BP&D high-school student Ss and with the 32 undergraduate Ss of the visual-mode pilot study, which was conducted in the same manner as that reported by BP&D. Mention has already been made of the higher 'failure' rate among the Reid Ss (31% vs 6% for 'pilot' Ss). When the Reid 'non-finishers' were assigned a TTC score of 128+, the

median TTC for the whole Reid sample was found to be 91.5; that for the pilot Ss was 63.5. In contrast to the BP&D 'failure' rate of 47%, and their median TTC of 116.5, the undergraduate Ss as a whole performed rather well, suggesting that better-educated persons are more easily made aware of syntactic pattern differences. However, the discrepancies between the two undergraduate samples with respect to both the above measures strongly indicate that the aural task occasioned greater difficulty. In addition, the difference between the TTC means for the two groups was significant at the 0.05 level (t = 2.37).

Paradoxically, the visual pilot Ss - the best performers of the three experiments - had the worst distribution of mean LET-ranks. Their concept-formation strategies were so variable that no significant statistical evidence of systematicity could be found. There seems to be an inverse correlation between ease of task (as measured by TTC) and consistency of CCF strategy. One might speculate that those Ss for whom the task was relatively easy were able to take advantage of intuitional 'leaps' and resolve more than one feature-distinction at a time. Although 'wholistic conservative focusing' is the most prevalent CCF strategy, intermittent 'focus gambling' is not uncommon (see Deese & Hulse, 1967). The net result of such strategy-shifting, where successful, would be to reduce TTC, but confound LET-patterns. In both these respects, the Reid Ss

were intermediate between the 'pilot' Ss and those of BP&D, lending further support to the conclusion that aural processing is more difficult.

<u>Analysis of Individual Performances</u>

It has been suggested that the S-sample drawn for this experiment included a disproportionate number, compared to the pilot study, of what might be called 'extreme' performers. The cause of this aberration lay in the experiment itself as a measuring procedure, rather than with random sampling as such; this matter is discussed further in Chapter Four. In order, however, to obtain a clearer picture of the internal composition of this subpopulation, two 'Principal Factors' factor analyses were performed, using the Ss as variables. The two basic data-arrays (as in Table 1) were transposed for input to the analysis, so that yes-response counts as a function of stimulus-category served as replications in the one case, and those according to error-category in the other. The Ss thus entered the analysis as a postulated swarm of points in 64-dimensional space. The correlations among individual S-performances, as indices of inter-point density, were used to compute a more parsimonious estimate of the dimensionality of the n-space, and the co-ordinate axes of this reduced space were taken to be the principal 'Factors'. The original variables were then expressed linearly in terms of their 'distance' from these orthogonal axes (Harman, 1967).

TABLE 8

SUMMARY OF FACTOR ANALYSES
OF SUBJECTS AS VARIABLES[1]

	Male	Stimulus-Factors[2]							Error-Factors[3]				Feml	Stimulus-Factors							Error Factors			
Tgt	Ss	1	2	3	4	5	6	7	1	2	3	4	Ss	1	2	3	4	5	6	7	1	2	3	4
	6¬		*							8			7					*			7	6		
ADF	11		6			*			*				13					*			*			
	21¬		6		5					7			20¬					*			7	7		
	25¬					7				*			27					*			8			
	3			*					8	6			2			*	5				*			
ADN	10			*					*				8			*					*			
	17				7					*			21¬				8					*		
	29			*					*				31			*					7	6		
	5						*		*				11						6		7	6		
AIF	9¬					5				*			15						*		*			
	23						*		*				22						*		*			
	26						*		*				30						7		7			
	2	*							8				5	*							8	5		
AIN	16	*							7	7			9	*							*			
	18	*							8	5			18¬									5	7	
	31	*							*				29	*							7	6		
	4		*						*				6		*						*			
PDF	14		*						*				16¬					7				8		
	24¬		*						*				24		*						*			
	32			6		5				6		5	26		*						*			
	8				8				7	7			3				*				*			
PDN	12¬		5								*		12				*				*			
	22¬		*								7	5	17				*				*			
	27				*				*				32¬			6						*		
	7¬	6									*		10¬							6			*	
PIF	13					5			*				14								7	5		
	19¬	6							5	5		5	23¬					7			8			
	30					5			8		5		25¬										8	
	1							*	*				1							*	*			
PIN	15	7								*			4¬	8								*		
	20							*	*				19							*	*			
	28¬	5								7			28¬	7								*		

[1]coefficients to nearest tenth, X 10; '*'= coeff. ≥ 0.9.
[2]Promax rotation of Princ. Factors for stimulus-type.
[3]Varimax rotation of Princ. Factors for error-type.
'¬' = Ss who did not achieve criterion.

Since the principal axis (Factor 1) is a linear function, computed to be a least-squares best fit to all points in the multidimensional swarm, any internal organization or clustering that might exist in the sample would tend to remain obscure. For analyses whose purpose is to estimate the number and membership of the major groups of intercorrelated variables (Ss, in this case), simpler factor structure can be obtained through rotation of the principal axes to a position where Factor 1 is the best-fitting line through the major cluster in the positive quadrant ('Varimax' method). Maximally simple factor structure can be computed through oblique rotation of the reference axes to positions where they are the centroids of successively less dense clusters; this, however, tends to destroy their orthogonality, and yields a solution where the factors are no longer independent.

The advantage of factor analysis is that it enables one to conflate a multitude of individual behaviors into an economical set of 'most prevalent' behavior-patterns, while at the same time quantifying the degree to which each individual exhibited each such behavior-pattern. The technique, however, extracts at most only (n-1) real factors, where n is the smaller dimension of the input matrix; thus, no more than seven meaningful clusters could possibly result from the present analyses. The analysis of 'yes-count per stimulus-type' performances defined seven

types of response-pattern, corresponding to the seven target-groups most similar in response-profile (see Figure 1). These seven factors were the result of oblique rotation ('Promax' method) of the principal axes to 'best group-fit'; Table 8 summarizes the 64 Ss' most significant behavior components with respect to these factors. Not surprisingly, the one response-pattern that failed to emerge as a factor was that for 'typical PIF response'. With five of the eight PIF Ss not achieving TTC, it was not possible for any coherent group-behavior as such to develop; correlations could not but be poor, among essentially random response-patterns. Three of these Ss, it will be noted, exhibited mainly AIN- or PIN-like behavior; they would seem to account largely for this group's high incidence of Modality and Voice+Modality errors (see Figure 1).

The PIN Ss were another aberrant group. On the average, they appeared to respond 'normally' (see Figure 1), but only half of them performed as expected. The other four, three of whom did not achieve criterion, responded almost indistinguishably from most AIN Ss, and seem to be responsible for the PINs having the highest mean for Voice-errors. Similarly, half of the male ADF subgroup did not form a complete concept, and responded in the manner typified by PDF Ss; they correlated rather poorly with each other and with their female counterparts, as did the male PDN Ss, to a lesser extent. In general, Ss who responded atypically, i.e., in a manner poorly correlated with the

others in their target-group, were in most (but not all) cases non-finishers. Likewise, most, but not all, of the twenty who did not achieve criterion performed unlike their group-mates. But these, as has already been pointed out, did not all respond randomly; 14 of them, in fact, had, by trial 128, formed two-thirds of their target-concept, and were erring in either the Voice or the Modality dimension, only.

The situation with respect to the 'yes-counts per error-category' data was somewhat different from the foregoing: where, in the analysis of response-pattern _per se_, the group-factor solution seemed the most appropriate description, the nature of the alternate data-arrangement forced a 'general factor' solution which even oblique rotation could not disperse. This was because most _S_s, whether they learned to differentiate target-sentences from other types or not, made a substantial contribution to the 'correct response' category (see column 1 of Table 4). The result of the direct Principal Factors solution on this data-set was a heavily-loaded first factor, accounting for 76% of the variance in the reduced common-factor space, and three rather lightly-loaded factors; together, the four factors accounted for 98% of the total variance.

Factor 1 obviously represented a 'No error' performance pattern, and a check of the individual error-patterns of the relevant _S_s verified this. However, many

whose large coefficients (0.9 or better) in Factor 1 classified them as typically error-free responders were not really so, having made a substantial number of Voice-errors, as well as correct yes-responses. Also, while the greatest loadings on Factors 2 and 3 were found to represent the Voice- and Modality-error contributions of 'non-error-free' Ss, these coefficients were not clear-cut enough to characterize the factors unequivocally. In other words, the unrotated Principal Axes did not fit the S-clusters as best they might.

The Varimax rotation of the axes seemed to produce the clearest and most appropriate classification of the Ss according to typical error-pattern; a subsequent oblique solution seemed only to becloud the picture once more, this time weighting the significant No error:Voice-error proportions in favor of the latter. The Varimax factor pattern is also summarized in Table 8. Factor 2 emerged clearly as the 'Voice-error' response pattern, and Factor 3 appeared to be that for Modality-errors, with some Voice+Modality-error contribution from several Ss. The weightiest contributions to Factor 4 (0.57 at most) were from Ss who made some Modality-errors, and some other type of error besides Voice+Modality, thus characterizing this factor as an 'everything else' behavior pattern. This analysis agreed highly with the results of the Tukey test on error-category means (Table 6), where very similar

clusterings of response-pattern were brought to light. In this factor analysis, too, the effects of not achieving criterion became apparent: eleven of the non-finishers were classified as definitely Voice-error prone, three as Modality-error committers, and the remaining six as inclined to a variety of error-types.

CHAPTER FOUR

DISCUSSION

The Experimental Technique

The experiment reported herein has demonstrated fairly conclusively the viability of the CCF technique as a means of assessing people's awareness of the syntactic properties of vocal messages. The principal aim of the study is thus realized. But the use of the oral-aural medium and other innovations have brought to light some problems; these deserve elaboration, and give rise to certain recommendations and *caveats*.

Sex as an Independent Variable

It would have been somewhat disconcerting, from the point of view of linguistic theory, if *Ss*' response-patterns had been significantly differentiated as a function of sex. Sex differences have been observed in language acquisition (Gesell, 1940, p. 199) and in linguistic studies involving sex-correlated cultural factors, such as honorifics and color-names. But there is no reason to suppose that normal

English-speaking adult males and females should differ in their awareness of syntax. It would have been equally surprising, of course, if there had been no discernible sex-effects at all. Without incorporating sex as a factor in the experimental design, however, there would simply have been no way of isolating variance components involving sex, nor of estimating their relative importance. This ancillary factor, had it been significant and uncontrolled in this experiment, would have tended to inflate the estimate of error-variance. Since sex-related effects have often been reported in studies of linguistic skill, it seems desirable, in general, that replications in a linguistic experiment always be divided equally between the sexes.

The Need for Complete Concept-formation

Mention has already been made of the difficulty of obtaining a homogeneous sample of *S*s, even from such a restricted subpopulation as undergraduates, and of the disturbing proportion of 'extreme' performers in the present sample. Twenty of these, it has been shown, were the unexpected result of limiting the number of trials to 128. It seems likely that at least a few of these non-finishers would have been discarded altogether, had this experiment been a TTC task: the six who were still erring in all three syntactic dimensions at the last trial might well never have solved the problem at all. Three *S*s, on the other hand,

made no contribution whatever to the analysis of errors, except to add to an already significant mean for correct yes-responses (see Table 6).

It is chiefly in the analysis of concept-formation strategies that these extreme cases made their presence felt. Knowledgeable Ss contributed heavily to the yes-response counts in the 'correct' category, but little or nothing by way of acquisition pattern. One suspects that a great deal of their concept-formation took place in the first four trials, thanks to a very explicit set of instructions, or soon thereafter, through their intelligent analysis of negative instances. But then, what of the 'random responders'? Their contribution to both yes-response counts and LET scores only added to the general error-variance; they were apparently unable to profit from explicit instruction as well as the others, or to derive the full informational content of every 'correct' or 'wrong' reinforcement. The result, as has been seen, was a pattern of LET means about which little can be said with any degree of certainty.

It comes as no surprise that people (undergraduate people, at least) vary widely in their ability to understand instructions or perceive syntactic features; nor is it unexpected that a random sample of Ss should include a few aberrant performers. Still, it seems highly improbable that some 14% of the population should be classed as deviant with

respect the linguistic skills in question: what seems more likely is that the experiment itself is an imperfect instrument for measuring these abilities. A less explicit set of instructions, for example, besides saving some experimental time, would have required abler Ss to go through the complete trial and error procedure which is the sine qua non of the CCF technique, thus contributing some information about relative subjective dominance relations among syntactic features. This would not necessarily have made random responders of less able Ss, for a great many of them, after a dozen years of schooling and a pre-experimental 'refresher course', still could not grasp the active/passive distinction, and said so. If distinct 'constructional meanings' (S(t); see pp. 21-2, above) for different syntactic patterns exist for a S, that difference is what he will respond to, not grammarians' type-labels.

The foregoing proposal does, however, increase the likelihood of longer trial and error periods with all Ss, requiring a more open-ended experimental procedure. The imposition of a fixed number of trials now seems to have been ill-advised, for it inhibited full concept-formation on the part of some Ss, and elicited a great deal of redundant information from several others. TTC scores themselves, as has been argued (see Chapter 1), remain of no consequence, as do the absolute numbers of yes-responses in any given category: it is rather the distribution of responses which matters, and percentage of TTC is more informative than raw

frequency-counts, in that respect. No form of numerical data is of value, it must be concluded, unless full concept-formation has taken place.

It is not proposed that the number of trials be completely unbounded, for S-fatigue will set inexorable limits on that. However, with shorter instruction time and a tightening up of sentence-presentation procedures, it ought to be possible to make Ss process a substantially greater number of sentences without exceeding those limits.

Time as a Covariant Measure

One of the most efficient statistical techniques for 'partialling out' the contaminating effects of individual differences is the analysis of covariance. The 'total time' measure observed in this study proved too gross to be of any utility, but the correlation that was found between time and errors (see Table 2) suggests further possibilities. The time variable was, in fact, found to correlate significantly with all four subtypes of response at the 0.05 level (ignoring the oldest S), positively with wrong 'yes' and 'no' responses, negatively with correct ones. This is apparently not unusual, and it generally indicates that those Ss who knew what they were doing responded quickly, and that those who did not took longer.

It must be pointed out, however, that a CCF task, if

fully implemented as suggested above, is one where no S knows what he is doing, initially, and every S does, at the end. In between, certainty varies as a function of stimulus and the S's current notion of target. This was certainly evident in the present experiment, where many Ss, once they had some idea of their target-type, hesitated noticeably before responding, when presented with a target-like stimulus sentence. Pauses were especially long before Voice-opposites of target, in the later trials.

A cumulated response-latency measure, if nothing else, would exclude the uninformative gaps between trials which inflated the present 'total time' measures. The individual response-latencies, on the other hand, might provide a more sensitive measure of type-feature perceivability than yes or no responses, in which there is bound to be some degree of guesswork. In any case, covariances between these dependent measures would be available for examination. The need for more efficient stimulus-presentation methods has already been pointed out; with the ever-increasing availability of computers for real-time control of experiments, the implementation of both innovations seems quite feasible.

Theoretical Considerations

The Primacy of Meaning

Considering the latitude of individual variation among Ss, there was a remarkable degree of correspondence between the results of this present study and that of BP&D. That agreement not only cut across lines of perceptual modality, but also across sex, age, and educational advancement. To be sure, the more literate Ss performed somewhat less predictably, but that, as has been shown, is partly the result of a more difficult stimulus-presentation mode. In general, the conclusions reached by BP&D have been corroborated and given a broader base of support. At the same time, the cross-modality differences found by Fodor & Garrett (1967) must be discounted: had the present CCF experiments also been conducted with a sample size of only ten, they might well have yielded similar conflicting results.

People in general are quite unaware of syntax when they converse. Paradoxically, the ways in which vocal symbols may be temporally sequenced are known to and used by every fluent speaker, but the syntax itself seems to have no autonomous existence, except as can be inferred from observed utterances. If this is true for grammarians, so much more so is it for naive language-users. Furthermore, people would appear to vary widely in their ability to infer

syntactic structure from utterances in the detached manner of the linguistic analyst, even when required to do so. This is attested by the great disparities in ITC observed in both experiments, and by the disturbing number of Ss to whom syntactic patterns were hopelessly opaque. The distribution of this talent in the population is, however, another matter, one which neither this nor the BP&D experiment was designed to probe.

What these studies have shown is that ordinary language-users can only with difficulty be made to divorce themselves from semantic considerations, when analyzing sentences. The dominance-hierarchies inferred among the three syntactic dimensions are quite the reverse of those predicted by either surface structure or transformational factors. Active and passive sentences, being superficially distinct in so many respects, should have been most easily distinguishable; yet they were not, except by the PDF group in the present study. Affirmation versus negation, on the other hand, differing the least, superficially, should have constituted the most difficult distinction, particularly in the aural mode of presentation. But this too, despite the higher incidence of Modality errors in the aural experiment, was not generally the case. Neither CH nor DTC can account for the observed data, and, as BP&D have pointed out, a theory of 'markedness' offers only a partial explanation.

The observed facts can be accounted for much more

adequately when one considers the S(c)/S(t) dichotomy (see pp. 21-4, above). It seems abundantly clear that the overwhelming majority of _S_s were quite oblivious to syntactic differences at the start, and remained so in varying degrees until the end of the experimental session. If this had not been the case, more than a small minority of them would have performed as one would expect grammarians to do, that is, resolve the most extensive pattern-difference - Voice - first. Pure syntax can therefore be eliminated from consideration, leaving meaning as the basis for people's apprehension of sentences. Of the two aspects of meaning postulated, furthermore, S(c) must be ruled out as a possible factor, because of the systematic variation in lexical burden that was built into the stimuli. Sentences uttered in such noncommunicative situations tend not to have much meaning (in the everyday sense) at any rate, as BP&D have indicated.

Most _S_s did eventually come to see the similarities among such apparently divergent sentences as

The cheese won't be eaten by mice.

and

John isn't being blamed by Bill.

This is a strong indication that they had become sensitive, if not to the common syntax itself, at least to something they could associate with it, _i.e._, S(t). They had been made aware, in other words, of that component of meaning

which even lexically distinct sentences can share, and in that respect, at least, can be said to have the same meaning. The general difficulty of the task attests to its unnaturalness: S(c) and S(t) are simultaneously conveyed by means of a never-empty syntactic 'vehicle' as an integral message, and there is normally no need for the meaning of the vehicle to be disengaged from that of its contents. This may account for the fact that S(t) eluded the attention of psycholinguistic experimenters for so long.

'Markedness' in Type-features

Once *Ss* had adverted to the necessity of discovering other classificatory criteria than content, tense, and the like, they were faced with a task analogous to a Gestaltist figure-ground puzzle. They had to learn to ignore that which they by lifelong habit attended to the most - S(c) - in order to focus on that which usually escapes their notice, S(t). Despite wide variation among individuals in attending to these normally covert notions, it seems clear from the data that 'activeness' and 'passiveness' are the most obscure, and the other features rather less so.

In view of the numerous arbitrary aspects of language, it is no surprise to find a lack of correlation between communicative import and syntactic prominence. Yet, while one can accept that people simply learn to attach a particular significance to a particular form, complex or

simple, and behave accordingly, one wonders at the associative disparities acquired by speakers of English. There have always been, it seems, the Mood and Modality features on the one hand, and the active/passive distinction, somehow different in kind, on the other. Thus, while the S(c)/S(t) distinction allows previously recalcitrant results to be accounted for (see BP&D), it does not, per se, explain why the Voice alternation has been so impervious to analysis.

People's awareness of activeness and passiveness may be context-dependent in a way which the other S(t) notions are not: unless there is a compelling need to distinguish the one 'focus' from the other, the distinction seems not to be made at all. Such need arises in the CCF experimental situation, when Ss have solved two 'problems' and continue to receive 'wrong' reinforcements. It also occurs in the context of discourse, where, for one reason or another, attention is to be drawn to the 'acted-upon' rather than the actor (or vice versa). Such is not the case with Mood and Modality features. To be sure, interrogation and negation can be distinguished from declaration and affirmation, should the need arise; but these latter four notions differ from the Voice-opposites in that they are self-evident within a given sentence, and seldom require the presence of contrast in order to be perceivable as S(t) components.

Mood and Modality 'color' an utterance in a fairly

obvious unitary fashion, where Voice, contrary to intuitive expectation, does not. Negation and interrogation, performatively, are a 'different manner of speaking', inherently 'marked', as opposed to 'neutral' or 'unmarked'. The data strongly suggest, however, that there is no inherently 'marked' Voice, as such: the difference in meaning is so subtle, surface differences notwithstanding, that some degree of analytic experience is required in order for Voice in utterances to be apparent at all.

The active/passive distinction was originally formalized as an optional syntactic transformation mainly because doing so made for a simpler and less *ad hoc* grammar (Chomsky, 1957). Like Zellig Harris before him, Chomsky was attempting to describe, in a maximally simple way, the very general but superficial correspondences among sentence-types observed in discourse (see Z. Harris, 1952, 1957). It now seems significant that, even in a meaning-independent framework, Chomsky should consider actives to be as derivable from passives as passives from actives, the only cogent deciding factor being economy of description (1957, pp. 79-80).

The first serious attempts at accounting for meaning in sentences immediately ran into difficulties with Voice. In order to devise a workable meaning-oriented model of grammar, Katz & Postal (1964) had first to deny Chomsky's 1957 claim (p. 101) that actives and passives have different

meanings, and hence must be related by means of an optional meaning-changing T-rule. But Katz & Postal were not about to give up the simplicity and generality of the passive transformation, and had then to postulate a semantically empty transformation-triggering base morpheme to account for the now 'stylistically variant' Voice-opposites (1964, p. 73).

The results of these CCF experiments suggest that the grammar theorists were in some respects correct, all along. Differences in active-passive syntax, being entirely intrasentential in manifestation, are quite adequately describable in a sentence-grammar. But what 'triggers' the actual communication of the one focus or the other is another matter. It seems to be fairly obvious to a listener, from the Mood and Modality manifestations in a sentence, that the speaker has either asserted or questioned something, in an affirmative or negative fashion. But, assuming a desire on the part of the speaker to draw attention to the 'acted-upon' rather than the actor, or vice versa, whether such focus is apprehended as a focus, through passive syntax alone, would seem to depend on extrasentential considerations. In a sense, actives sometimes mean the same as passives, and sometimes do not, for the sentential meanings assumed to reside in distinct S(t) components can be distinguished only under exceptional circumstances, and sometimes not at all.

This does not imply that there are instances where speakers' intended meanings fail to be conveyed, for that is obviously not the case. "The car has been driven by Susan" is understood as meaning just what it says, in or out of discourse. But the result of the present study lead one to believe that "Susan has driven the car" would be equally communicative, since S(c) - the lexical items and their actor-object and other semantic relations - is the same in both. Whichever S(t) is conveyed, it is an integral part of the whole meaning of the sentence, and achieves no prominence of its own except through explicit contrast.

The active/passive dichotomy, in sum, appears to be entirely different in kind from Mood and Modality,and this is where years of linguistic tradition must be gainsaid. Every sentence must indeed have a Voice, and the only options available in English are active and passive. But it is not every sentence that <u>can</u>, let alone must, be either active or passive: only S(c) constellations which include a transitive verb can be expressed in the passive. This would account at least in part for the relative rarity of this form of locution, and its concomitant unfamiliarity (Goldman-Eisler & Cohen, 1970). It is precisely because of this restricted applicability that Passive Voice has been regarded as a 'marked' state of affairs, a departure from the usual, and formalized as an option in grammars. This CCF experiment and its antecedent have shown, however, that passiveness can be considered 'marked' only in the formal

sense, for it is highly imperceptible to the ordinary language-user.

<u>Implications for Linguistic Theory</u>

The early studies of sentential complexity, like early Chomskyan theory, tended to oversimplify. Matters of peripheral concern, such as nonlexical meaning, presuppositions, and the like, were not merely put aside for the sake of limiting the scope of an investigation; they were often dismissed as totally irrelevant, or overlooked altogether. Man, despite centuries of bedeviling complexity, was assumed to use language as simply as the most powerful formal rewriting system could describe (Chomsky, 1959, p. 56). But the fact is, and always has been, that language-use is 'compositional' in nature, exactly as Katz & Fodor envisioned it (1963, pp. 482ff.): utterances are endowed with a multiplicity of meanings in a multitude of ways.

A recent proposal for a better formal treatment of passivization and other syntactic processes (Chambers, 1970) is founded on the traditional assumption that Passive Voice is 'marked', while the Active is not. Empirical evidence is cited in support of this intuition, where a speaker's selection of Voice is shown to be conditioned by the "encoding situation" (pp. 4-6). The assumption is a fruitful one, for it enables a number of grammatical

processes, including passivization, to be described in a fairly general way, through the introduction in deep structure of the obligatory but mobile semantic 'primitives' [±FOCUS] and [±TOPIC]. The present study and that of BP&D offer further evidence that the selection and apprehension of Voice is in some way dependent on the situational (and discourse) context, but must deny, however, that either of the two Voice alternants can be considered 'marked' in the same sense as can Mood, Modality, and certain other features.

The formal semantic feature approaches to language description, like these CCF experiments, are an attempt to grapple with an elusive set of abstract notions, in order to discover and delineate whatever regularities there might be. Where the enterprises differ, of course, is that the former ascribe those regularities to sentences, where the latter prefer to draw conclusions about people. With such divergent goals, it is doubtful that this or any other psycholinguistic finding will have much impact on current 'competence' theories. Most grammarians are, by their own stipulation, in the business of describing language, not speech, and their formalizations do not include variables from the empirical domain. Even if they did, Chambers points out, "current generative grammar could not in principle deal with the discourse context, since it has as its purview the sentence itself and not its context" (p. 6).

Furthermore, the aspect of linguistic ability reflected in the present findings represents the _typical_ speaker-hearer, not the 'ideal' one; Chomskyans, therefore, are not likely to complicate their descriptions just to account for a formally inconsequential notion.

Fillmore's 'case' grammar (Fillmore, 1968) appears to be a more viable theory of sentence-description. This model takes explicit account of 'actor' and 'acted-upon' relationships, among others, and uses these semantic notions to generate fairly large classes of related sentence-types in relatively simple ways. Since people appear to be generally unaware of active or passive focus itself, as long as the 'underlying' actor-object relationships are undisturbed, case grammar would seem already to incorporate some of the inferences that have been made here. But it, too, is only a 'competence' grammar, a derivative of the standard Chomskyan model, and formally unequipped to deal with either psychological reality or discourse factors.

What is needed is a revised theory of linguistic competence. As Chafe puts it,

> A theory of competence must bear a relationship to language use, and there is no reason to assume that it is a virtue for this relationship to be as obscure as possible . . . To search for points of contact between performance and the syntacticist model is a very unsatisfying pursuit. (1970, p. 66)

Chafe's model seems the most promising to date. He takes

sentence *production* as his goal, not mere formal sentence 'generation', and begins with the formation of a semantic structure, taking the total utterance-context into account. His view of Voice, in particular, is in accord with the results and conclusions reported here. Active and passive locutions are both fundamentally unmarked; the latter provides a means of preserving the usual 'left to right' order of 'old' and 'new' information in a sentence, when the 'patient' NP can be presupposed to be familiar to the hearer.

Besides admitting that age-old intuitions about language use may, after all, be correct, a meaning-based grammar such as Chafe's is a rich potential source of empirically testable hypotheses. Only such a formulation can hope to reflect the 'internalized knowledge' people draw upon when they converse, and, at the same time, form the core of a viable theory of linguistic performance. The Chomskyan 'revolution' has revealed the ultimate sterility of describing linguistic products apart from their producers, but it has also taught analysts to sharpen their thinking, and has provided them with better descriptive tools. Sooner or later, these tools will have to be applied to the facts of language use, and a more realistic theory of linguistic competence will be the result. It is hoped that the present work will find a place in it.

CHAPTER FIVE

SUMMARY AND CONCLUSIONS

The psychological complexity of sentences and its relation to generative-transformational grammar theory has been subjected to experimental testing for a number of years, with generally equivocal results. Critical examination of the relevant literature shows many past studies to have been theoretically, methodologically, and foundationally inadequate. A three-dimensional conjunctive view of sentential type-features proposed by Baker, Prideaux, & Derwing ('BP&D', in press) led these authors to apply the conjunctive concept-formation paradigm to the assessment of the relative subjective prominence of those feature-dimensions. Their finding of systematic response-pattern differences between Mood (declarative or interrogative) and Modality (affirmative or negative) features on the one hand, and Voice (active or passive focus) on the other, suggested that Ss respond to the *meaning* of a syntactic pattern, not the manifest pattern itself, and that the Voice alternatives are different in

kind from the other 'features'.

An experiment was performed, in which the BP&D investigation was repeated, using aural rather than visual stimulus-presentation, and a fixed number of trials, with more sophisticated Ss; these were subgrouped with respect to sex within a 2 X 8 X 2 X 2 X 2 factorial design (partially repeated measures). A set of 128 sentences, systematically varied with respect to Voice, Mood, Modality, tense, and lexical content, was presented, in one of eight random orders, to each of 32 male and 32 female undergraduate Ss. Their task was to infer, from differential correct/wrong reinforcement of yes/no judgments, a sentence-type characterized by one of the eight possible combinations of Voice, Mood, and Modality features, and to respond 'yes' to all and only such sentences. The other variations in sentence form were uncorrelated with these variables. Yes-response counts, as a function of randomly-assigned target-type and stimulus-sentence type, were analyzed as evidence of advertence to sentence-types thought - rightly or wrongly - to have been target-type. The last trials in which sentences differing from target in Voice or Mood or Modality received a (wrong) yes-response were interpreted as indices of the relative perceptibility of these syntactic properties.

An analysis of variance showed the bulk of yes-responses to have been correct ones, to target-type

sentences ($p < 0.01$); next in predominance were those to Voice-opposites of target-type, that mean differing significantly from all others ($p < 0.01$). The proportions of yes-response as a function of target and type of error correlated highly with the corresponding values observed by BP&D ($r = 0.90$). 'Last error trial' means were rather less systematic, due to failure on the part of 20 Ss to form a stable three-way concept of target-type by trial 128. Factor analysis of the individual S-performances outlined four main patterns, with most Ss exhibiting mainly 'correct response' and 'Voice error' behavior, and 20-odd 'deviant' Ss classed as generally error-prone.

A comparison of the median 'trials-to-criterion' for this study with those of the BP&D high-school Ss and the undergraduates in a visual mode pilot study suggested that the aural task is, in general, more difficult. The resultant inability of 31% of the Ss to form a full concept within 128 trials underlined the advantage of a necessarily arbitrary criterion over a fixed number of trials. The feasibility of aural stimulation in such an experiment was nevertheless judged to have been demonstrated, and the BP&D conclusions given a broader base of support. Whereas negation and interrogation, as components of sentential meaning, are relatively obvious and 'marked', with respect to affirmation and declaration, active and passive Voice ('focus') *per se* are not as readily distinguished. It can thus justifiably be assumed that every sentence must

manifest one of the Mood and one of the Modality features; its Voice/focus, however, would seem to depend on extrasentential context factors, and is constrained by the transitivity of the main verb, in any case. While these conclusions are of potential utility to a performance-oriented theory of linguistic competence, they would seem to carry little import for current formal theories of language description, which exclude empirical variables and extrasentential context from consideration.

REFERENCES

Baker, W. J., Prideaux, G. D., & Derwing, B. L. Grammatical Properties of Sentences as a Basis for Concept Formation. _Journal of Psycholinguistic Research_, in press.

Bever, T. G. A Survey of Some Recent Work in Psycholinguistics. In W. J. Plath, _Specification and Utilization of a Transformational Grammar_. Yorktown Heights, N.Y.: I.B.M. Corp., 1968.

Bever, T. G., Lackner, J. R., & Kirk, R. The Underlying Structures of Sentences Are the Primary Units of Immediate Speech Processing. _Perception and Psychophysics_, 1969, 5, (4), 225-34.

Bloomfield, L. _Language_. New York: Holt, 1933.

Chafe, W. L. _Meaning and the Structure of Language_. University of Chicago Press, 1970.

Chambers, J. K. Focused Noun Phrases in English Syntax. Unpublished doctoral dissertation, University of Alberta, 1970.

Chomsky, N. _Syntactic Structures_. The Hague: Mouton, 1957.

Chomsky, N. Review of B. F. Skinner, _Verbal Behavior_. _Language_, 1959, 35, 26-58.

Chomsky, N. _Current Issues in Linguistic Theory_. The Hague: Mouton, 1964.

Chomsky, N. _Aspects of the Theory of Syntax_. Cambridge, Mass.: M.I.T. Press, 1965.

Clifton, C., Kurcz, I., & Jenkins, J. J. Grammatical Relations as Determinants of Sentence Similarity. _Journal of Verbal Learning and Verbal Behavior_, 1965, 4, 112-7.

Clifton, C., & Odom, P. Similarity Relations among Certain English Sentence Constructions. _Psychological Monographs_, 1966, 80, 5, 613.

Deese, J., & Hulse, S. H. _The Psychology of Learning_ (3rd ed.). New York: McGraw-Hill, 1967.

Derwing, B. L. Transformational Grammar as a Theory of Language Acquisition. Cambridge University Press, in press.

Ervin-Tripp, S. M., & Slobin, D. I. Psycholinguistics. Annual Review of Psychology, 1966, 17, 435-74.

Fillenbaum, S. On the Use of Memorial Techniques to Assess Syntactic Structures. Psychological Bulletin, 1970, 73, 231-7.

Fillenbaum, S. Psycholinguistics. Annual Review of Psychology, 1971, 22, 251-308.

Fillmore, C. J. The Case for Case. In Bach, E., & Harms, R. T. (Eds.), Universals in Linguistic Theory. New York: Holt, Rinehart, & Winston, 1968, 1-88.

Fodor, J. A., & Bever, T. G. The Psychological Reality of Linguistic Segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-20.

Fodor, J. A., @ Garrett, M. Some Reflections on Competence and Performance. In Lyons & Wales, 1966, 135-54.

Fodor, J. A., & Garrett, M. Some Syntactic Determinants of Sentential Complexity. Perception and Psychophysics, 1967, 2, 289-96.

Fodor, J. A., Garrett, M., & Bever, T. G. Some Syntactic Determinants of Sentential Complexity II: Verb Structure. Perception and Psychophysics, 1968, 3, 453-61.

Fodor, J. A., & Katz, J. J. (Eds.) The Structure of Language. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Garrett, M., Bever, T. G., & Fodor, J. A. The Active Use of Grammar in Speech Perception. Perception and Psychophysics, 1966, 1, 30-2.

Gesell, A. The First Five Years of Life. New York: Harper & Bros., 1940.

Goldman-Eisler, F., & Cohen, M. Is N, P, and PN Difficulty a Valid Criterion of Transformational Operations? Journal of Verbal Learning and Verbal Behavior, 1970, 9, 161-6.

Gough, P. B. Grammatical Transformations and Speed of Understanding. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 107-11.

Gough, P. B. The Verification of Sentences: the Effects of Delay of Evidence and Sentence Length. _Journal of Verbal Learning and Verbal Behavior_, 1966, 5, 492-6.

Gough, P. B. Experimental Psycholinguistics. In W. C. Dingwall (Ed.), _A Survey of Linguistic Science_. University of Maryland, 1971.

Harman, H. H. _Modern Factor Analysis_ (2nd ed., rev.). University of Chicago Press, 1967.

Harris, P. R. On the Interpretation of Generative Grammars. Unpublished M.Sc. thesis, University of Alberta, 1970.

Harris, Z. S. Discourse Analysis. _Language_, 1952, 28, 1-30.

Harris, Z. S. Co-occurrence and Transformation in Linguistic Structure. _Language_, 1957, 33, 283-340.

Hayes, J. R. (Ed.) _Cognition and the Development of Language_. New York: Wiley, 1970.

Hays, W. L. _Statistics for Psychologists_. New York: Holt, Rinehart, & Winston, 1963.

Hunt, E. B. _Concept Learning_. New York: Wiley, 1962.

Jacobs, R. A., & Rosenbaum, P. S. _English Transformational Grammar_. Waltham, Mass.: Blaisdell, 1968.

Johnson, N. F. The Psychological Reality of Phrase-structure Rules. _Journal of Verbal Learning and Verbal Behavior_, 1965, 4, 469-75.

Katz, J. J., & Fodor, J. A. The Structure of a Semantic Theory. _Language_, 1963, 39, 170-210.

Katz, J. J., & Postal, P. _An Integrated Theory of Linguistic Descriptions_. Cambridge, Mass.: M.I.T. Press, 1964.

Leech, G. N. Some Assumptions in the Metatheory of Linguistics. _Linguistics_, 1968, 39, 87-102.

Lyons, J., & Wales, R. J. (Eds.) _Psycholinguistics Papers_. Edinburgh University Press, 1966.

Marks, L. E., & Miller, G. A. The Role of Semantic and Syntactic Constraints in the Memorization of English Sentences. _Journal of Verbal Learning and Verbal Behavior_, 1964, 3, 1-5.

Martin, E. Toward an Analysis of Subjective Phrase Structure. Psychological Bulletin, 1970, 74, 153-66.

Martin, E., & Roberts, K. H. Grammatical Factors in Sentence Retention. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 211-8.

Martin, E., Roberts, K. H., & Collins, A. M. Short-term Memory for Sentences. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 560-6.

Matthews, W. A. Transformational Complexity and Short-term Recall. Language and Speech, 1968, 11, 120-8.

McMahon, L. E. Grammatical Analysis as Part of Understanding a Sentence. Unpublished doctoral dissertation, Harvard University, 1963. Reported in Miller & McKean, 1964, and Fodor & Garrett, 1966.

Mehler, J. Some Effects of Grammatical Transformation on the Recall of English Sentences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 346-51.

Mehler, J., & Miller, G. A. Retroactive Interference in the Recall of Simple Sentences. British Journal of Psychology, 1964, 55, 295-301.

Miller, G. A. The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. Psychological Review, 1956, 63, 81-97.

Miller, G. A. Some Psychological Studies of Grammar. American Psychologist, 1962, 17, 748-62.

Miller, G. A., & Chomsky, N. Finitary Models of Language Users. In Luce, R. D., Bush, R. R., & Galanter, E. (Eds.), Handbook of Mathematical Psychology, Vol. 2. New York: Wiley, 1963, 419-91.

Miller, G. A., & Isard, S. Some Perceptual Consequences of Linguistic Rules. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 217-28.

Miller, G. A., & Isard, S. Free Recall of Self-embedded English Sentences. Information and Control, 1964, 7, 292-303.

Miller, G. A., & McKean, K. O. A Chronometric Study of some Relations between Sentences. Quarterly Journal of Experimental Psychology, 1964, 297-308.

Miller, G. A., & Selfridge, J. A. Verbal Context and the Recall of Meaningful Material. American Journal of Psychology, 1953, 63, 176-85.

Perfetti, C. A. Lexical Density and Phrase-structure Depth as Variables in Sentence Retention. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 719-24.

Reber, A. S., & Anderson, J. R. The Perception of Clicks in Linguistic and Nonlinguistic Messages. Perception and Psychophysics, 1970, 8, 81-9.

Reid, J. R. SYNTASK: A Set of Programs for a Psycholinguistic Experiment. University of Alberta Computing Review, 1972, 5, 38-47.

Rohrman, N. L. The Role of Syntactic Structure in the Recall of English Nominalizations. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 904-12.

Savin, H. B., & Perchonock, E. Grammatical Structure and the Immediate Recall of English Sentences. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 348-53.

Schlesinger, I. M. Sentence Structure and the Reading Process. The Hague: Mouton, 1968.

Slobin, D. I. Grammatical Transformations and Sentence Comprehension in Childhood and Adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-27.

Slobin, D. I. Recall of Full and Truncated Passives in Connected Discourse. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 876-81.

Stuart, C. I. J. M. On the Empirical Foundations of Linguistic Description. Actes du Xe congres international de linguistes, 1967, I, 393-9.

Watt, W. C. (1970a) Comments on the Brown & Hanlon paper. In Hayes, 1970, 55-78.

Watt, W. C. (1970b) On Two Hypotheses Concerning Psycholinguistics. In Hayes, 1970, 137-220.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Yngve, V. H. A Model and an Hypothesis for Language Structure. Proceedings of the American Philosophical Society, 1960, 104, 444-66.

APPENDIX A

Structural Tabulation of NPs Employed in Stimulus-Sentences

The structures given are the simplified feature-matrices underlying the actual NPs used, as they might appear in a surface-structure Phrase-marker. The numbers given with the instances refer to the content-set(s) in which they appear.

Type	Surface Structure	NPs occurring as Agent (NP^1)	NPs occurring as Object (NP^2)
1a.	[+N] [-PRON] [+PROPER]	Bill (03) Mary (07) Paul (12) Susan (16)	John (03)
1b.	[+N] [-PRON] [Z] [+COMMON] [+PLURAL]	thieves (04) mice (11) dogs (15)	snowballs (06)
2a.	[+ART] [+N] [+DEF] [-PRON] [-DEM] [+COMMON] [The] [-PLURAL]	the boy (01, 13) the teacher (08)	the girl (01) the doctor (09) the ceremony (10) the cheese (11) the car (16)
2b.	[+ART] [+N] [-DEF] [-PRON] [-DEM] [+PLURAL] [a/many] [+COMMON]	a hunter (02) a clergyman (10) a plumber (14) many children (06)	an assignment (5)
2c.	[+ART] [+N] [+DEF] [-PRON] [+DEM] [+COMMON] [±PLURAL] [±PLURAL]	this student (05)	that grizzly (02) those papers (08)
2d.	[+N] [+N] [+PRON] [-PRON] [+POSS] [+COMMON] [±PLURAL] [±PLURAL]	our neighbors (9)	our house (04) our guests (12) my guitar (13) our drain (14) his cat (15)
3.	[+ART][+ADJ][+N] [-DEF][ ][-PRON] [-DEM][choc-][-PLURAL] [a][olate][cake]		a chocolate cake (07)

APPENDIX B

Systematic Distribution of Tense among Stimulus-sentences

PR = present-progressive PA = past-progressive
PE = perfect FU = future

Content sets	Sentence-types							
	1ADF	2ADN	3AIF	4AIN	5PDF	6PDN	7PIF	8PIN
01	PR	PE	PA	FU	FU	PA	PE	PR
02	PA	FU	PE	PR	PR	PE	FU	PA
03	FU	PA	PR	PE	PE	PR	PA	FU
04	PE	PR	FU	PA	PA	FU	PR	PE
05	PR	PE	PA	FU	FU	PA	PE	PR
06	FU	PA	PR	PE	PE	PR	PA	FU
07	PE	PR	FU	PA	PA	FU	PR	PE
08	PA	FU	PE	PR	PR	PE	FU	PA
09	PR	PE	PA	FU	FU	PA	PE	PR
10	FU	PA	PR	PE	PE	PR	PA	FU
11	PE	PR	FU	PA	PA	FU	PR	PE
12	PA	FU	PE	PR	PR	PE	FU	PA
13	FU	PA	PR	PE	PE	PR	PA	FU
14	PE	PR	FU	PA	PA	FU	PR	PE
15	PA	FU	PE	PR	PR	PE	FU	PA
16	PR	PE	PA	FU	FU	PA	PE	PR

APPENDIX C

The Stimulus-sentences, in Basic Presentation Order

These sentences were composed by Drs. Wm. J. Baker, Gary D. Prideaux, and Bruce L. Derwing, of the Linguistics Dept., University of Alberta, in January, 1971. They are displayed here in the basic randomized presentation-order devised for the Reid experiment, in blocks of 8 and 16.

The sentences were presented in cyclic fashion, once each. The first sentence presented to any S was the initial one of the 16-block in which the first occurrence of his target-type is, by design, the fourth sentence. The entry-points for the eight target-conditions are labeled 'TC1', 'TC2', . . . 'TC8'. After entry, the sentences were presented in numerical order, number 1 following number 128, until every sentence had been presented once. The eight presentation orders were re-recorded from the master tape, for use in the experiment.

Each sentence is uniquely coded. The first four characters of the code specify syntactic type, e.g., '4AIN', which indicates a type 4 sentence (Active Interrogative Negative). The last two characters specify lexical content-set, e.g., '08', which indicates content-set 8, "the teacher . . . grade . . . those papers". The 6th and 7th characters specify tense, e.g., 'PR', which indicates a sentence in present-progressive tense.

4AIN/PR08 ISNT THE TEACHER GRADING THOSE PAPERS? [TC1]
5PDF/FU09 THE DOCTOR WILL BE PHONED BY OUR NEIGHBOURS.
8PIN/PA12 WEREN'T OUR GUESTS BEING ENTERTAINED BY PAUL?
1ADF/FU13 THE BOY WILL PLAY MY GUITAR.
7PIF/PA03 WAS JOHN BEING BLAMED BY BILL?
6PDN/PE02 THAT GRIZZLY HASN'T BEEN TRACKED BY A HUNTER.
2ADN/PR14 A PLUMBER ISN'T FIXING OUR DRAIN.
3AIF/PE15 HAVE DOGS CHASED HIS CAT?

4AIN/FU16 WON'T SUSAN DRIVE THE CAR?
7PIF/PR11 IS THE CHEESE BEING EATEN BY MICE?
8PIN/PE04 HASN'T OUR HOUSE BEEN BROKEN INTO BY THIEVES?
1ADF/PR05 THIS STUDENT IS READING AN ASSIGNMENT.
3AIF/FU07 WILL MARY BAKE A CHOCOLATE CAKE?
6PDN/PR10 THE CEREMONY ISN'T BEING PERFORMED BY A CLERGYMAN.
5PDF/FU01 THE GIRL WILL BE HELPED BY THE BOY.
2ADN/PA06 MANY CHILDREN WEREN'T THROWING SNOWBALLS.

3AIF/FU14 WILL A PLUMBER FIX OUR DRAIN? [TC2]
6PDN/PA01 THE GIRL WASN'T BEING HELPED BY THE BOY.
7PIF/FU02 WILL THAT GRIZZLY BE TRACKED BY A HUNTER?
2ADN/PA13 THE BOY WASN'T PLAYING MY GUITAR.
4AIN/PA07 WASN'T MARY BAKING A CHOCOLATE CAKE?
1ADF/PA12 PAUL WAS ENTERTAINING OUR GUESTS.
8PIN/FU03 WON'T JOHN BE BLAMED BY BILL?
5PDF/FU16 THE CAR WILL BE DRIVEN BY SUSAN.

4AIN/PR15 AREN'T DOGS CHASING HIS CAT?
1ADF/PE04 THIEVES HAVE BROKEN INTO OUR HOUSE.
3AIF/PR06 ARE MANY CHILDREN THROWING SNOWBALLS?
8PIN/PE11 HASN'T THE CHEESE BEEN EATEN BY MICE?
6PDN/PA09 THE DOCTOR WASN'T BEING PHONED BY OUR NEIGHBOURS.
2ADN/PE05 THIS STUDENT HASN'T READ AN ASSIGNMENT.
7PIF/PA10 WAS THE CEREMONY BEING PERFORMED BY A CLERGYMAN?
5PDF/PR08 THOSE PAPERS ARE BEING GRADED BY THE TEACHER.

2ADN/PR11 MICE AREN'T EATING THE CHEESE. [TC3]
7PIF/PE16 HAS THE CAR BEEN DRIVEN BY SUSAN?
6PDN/FU07 A CHOCOLATE CAKE WON'T BE BAKED BY MARY.
3AIF/FU04 WILL THIEVES BREAK INTO OUR HOUSE?
1ADF/FU10 A CLERGYMAN WILL PERFORM THE CEREMONY.
5PDF/PE06 SNOWBALLS HAVE BEEN THROWN BY MANY CHILDREN.
4AIN/PE13 HASN'T THE BOY PLAYED MY GUITAR?
8PIN/PR09 ISN'T THE DOCTOR BEING PHONED BY OUR NEIGHBOURS?

2ADN/PA03 BILL WASN'T BLAMING JOHN.
3AIF/PE12 HAS PAUL ENTERTAINED OUR GUESTS?
6PDN/PE15 HIS CAT HASN'T BEEN CHASED BY DOGS.
1ADF/PA02 A HUNTER WAS TRACKING THAT GRIZZLY.
7PIF/FU08 WILL THOSE PAPERS BE GRADED BY THE TEACHER?
8PIN/PR01 ISN'T THE GIRL BEING HELPED BY THE BOY?
5PDF/PA14 OUR DRAIN WAS BEING FIXED BY A PLUMBER.
4AIN/FU05 WON'T THIS STUDENT READ AN ASSIGNMENT?

1ADF/PR09 OUR NEIGHBOURS ARE PHONING THE DOCTOR. [TC4]
8PIN/PA08 WEREN'T THOSE PAPERS BEING GRADED BY THE TEACHER?
5PDF/FU05 AN ASSIGNMENT WILL BE READ BY THIS STUDENT.
4AIN/PA04 WEREN'T THIEVES BREAKING INTO OUR HOUSE?
6PDN/PR06 SNOWBALLS AREN'T BEING THROWN BY MANY CHILDREN.
3AIF/FU11 WILL MICE EAT THE CHEESE?
2ADN/PA10 A CLERGYMAN WASN'T PERFORMING THE CEREMONY.
7PIF/FU15 WILL HIS CAT BE CHASED BY DOGS?

8PIN/PR16 ISN'T THE CAR BEING DRIVEN BY SUSAN?
5PDF/PE13 MY GUITAR HAS BEEN PLAYED BY THE BOY.
7PIF/PR07 IS A CHOCOLATE CAKE BEING BAKED BY MARY?
1ADF/PR01 THE BOY IS HELPING THE GIRL.
2ADN/FU02 A HUNTER WON'T TRACK THAT GRIZZLY.
6PDN/FU14 OUR DRAIN WON'T BE FIXED BY A PLUMBER.
3AIF/PR03 IS BILL BLAMING JOHN?
4AIN/PR12 ISN'T PAUL ENTERTAINING OUR GUESTS?

8PIN/PE14 HASN'T OUR DRAIN BEEN FIXED BY A PLUMBER? [TC5]
1ADF/PA15 DOGS WERE CHASING HIS CAT.
4AIN/PR02 ISN'T A HUNTER TRACKING THAT GRIZZLY?
5PDF/PA11 THE CHEESE WAS BEING EATEN BY MICE.
7PIF/PE05 HAS AN ASSIGNMENT BEEN READ BY THIS STUDENT?
6PDN/FU04 OUR HOUSE WON'T BE BROKEN INTO BY THIEVES.
2ADN/FU08 THE TEACHER WON'T GRADE THOSE PAPERS.
3AIF/PA09 WERE OUR NEIGHBOURS PHONING THE DOCTOR?

4AIN/PE10 HASN'T A CLERGYMAN PERFORMED THE CEREMONY?
6PDN/PE12 OUR GUESTS HAVEN'T BEEN ENTERTAINED BY PAUL.
5PDF/PE03 JOHN HAS BEEN BLAMED BY BILL.
3AIF/PA01 WAS THE BOY HELPING THE GIRL?
7PIF/PA13 WAS MY GUITAR BEING PLAYED BY THE BOY?
2ADN/PE16 SUSAN HASN'T DRIVEN THE CAR.
1ADF/PE07 MARY HAS BAKED A CHOCOLATE CAKE.
8PIN/FU06 WON'T SNOWBALLS BE THROWN BY MANY CHILDREN?

7PIF/FU12 WILL OUR GUESTS BE ENTERTAINED BY PAUL? [TC6]
2ADN/PR07 MARY ISN'T BAKING A CHOCOLATE CAKE.
3AIF/PE08 HAS THE TEACHER GRADED THOSE PAPERS?
6PDN/FU11 THE CHEESE WON'T BE EATEN BY MICE.
1ADF/PE14 A PLUMBER HAS FIXED OUR DRAIN.
5PDF/PR02 THAT GRIZZLY IS BEING TRACKED BY A HUNTER.
4AIN/FU01 WON'T THE BOY HELP THE GIRL?
8PIN/FU13 WON'T MY GUITAR BE PLAYED BY THE BOY?

1ADF/FU06 MANY CHILDREN WILL THROW SNOWBALLS.
3AIF/PA16 WAS SUSAN DRIVING THE CAR?
5PDF/PE10 THE CEREMONY HAS BEEN PERFORMED BY A CLERGYMAN.
2ADN/FU15 DOGS WON'T CHASE HIS CAT.
6PDN/PR03 JOHN ISN'T BEING BLAMED BY BILL.
8PIN/PR05 ISN'T AN ASSIGNMENT BEING READ BY THIS STUDENT?
4AIN/FU09 WON'T OUR NEIGHBOURS PHONE THE DOCTOR?
7PIF/PR04 IS OUR HOUSE BEING BROKEN INTO BY THIEVES?

6PDN/PR13 MY GUITAR ISN'T BEING PLAYED BY THE BOY. [TC7]
3AIF/PE02 HAS A HUNTER TRACKED THAT GRIZZLY?
2ADN/PE01 THE BOY HASN'T HELPED THE GIRL.
7PIF/PR14 IS OUR DRAIN BEING FIXED BY A PLUMBER?
5PDF/PR12 OUR GUESTS ARE BEING ENTERTAINED BY PAUL.
1ADF/PA08 THE TEACHER WAS GRADING THOSE PAPERS.
8PIN/PA15 WASN'T HIS CAT BEING CHASED BY DOGS?
4AIN/PE03 HASN'T BILL BLAMED JOHN?

1ADF/PR16 SUSAN IS DRIVING THE CAR.
5PDF/PA04 OUR HOUSE WAS BEING BROKEN INTO BY THIEVES.
3AIF/PR10 IS A CLERGYMAN PERFORMING THE CEREMONY?
8PIN/PE07 HASN'T A CHOCOLATE CAKE BEEN BAKED BY MARY?
2ADN/PE09 OUR NEIGHBOURS HAVEN'T PHONED THE DOCTOR.
7PIF/PA06 WERE SNOWBALLS BEING THROWN BY MANY CHILDREN?
4AIN/PA11 WEREN'T MICE EATING THE CHEESE?
6PDN/PA05 AN ASSIGNMENT WASN'T BEING READ BY THIS STUDENT.

5PDF/PA07 A CHOCOLATE CAKE WAS BEING BAKED BY MARY. [TC8]
4AIN/PA14 WASN'T A PLUMBER FIXING OUR DRAIN?
1ADF/FU03 BILL WILL BLAME JOHN.
8PIN/FU10 WON'T THE CEREMONY BE PERFORMED BY A CLERGYMAN?
7PIF/PE09 HAS THE DOCTOR BEEN PHONED BY OUR NEIGHBOURS?
6PDN/PE08 THOSE PAPERS HAVEN'T BEEN GRADED BY THE TEACHER.
3AIF/PA05 WAS THIS STUDENT READING AN ASSIGNMENT?
2ADN/FU12 PAUL WON'T ENTERTAIN OUR GUESTS.

4AIN/PE06 HAVEN'T MANY CHILDREN THROWN SNOWBALLS?
6PDN/PA16 THE CAR WASN'T BEING DRIVEN BY SUSAN.
8PIN/PA02 WASN'T THAT GRIZZLY BEING TRACKED BY A HUNTER?
5PDF/PR15 HIS CAT IS BEING CHASED BY DOGS.
7PIF/PE01 HAS THE GIRL BEEN HELPED BY THE BOY?
2ADN/PR04 THIEVES AREN'T BREAKING INTO OUR HOUSE.
3AIF/PR13 IS THE BOY PLAYING MY GUITAR?
1ADF/PE11 MICE HAVE EATEN THE CHEESE.

APPENDIX D

Instructions to Subjects

"Thank you for volunteering your help.

The purpose of this experiment is to study how people perceive similarities and differences among English sentences. You will hear 128 different sentences, in random order. No two of these sentences will be exactly the same, but many of them will be similar to each other in some way. Those sentences which are similar can be thought of as being of the same type, even though they are not identical to each other. Although you will hear many different sentences, there will be only a few different types of them; we will want you to concentrate on one type of sentence.

Sentences can resemble each other in many different ways. For example, a group of sentences could be all about the same topic, such as the following:

Jim is washing the dishes.
Have the dishes been washed by Jim?
The dishes weren't being washed by Jim.
Won't Jim wash the dishes?

Or, a group of sentences could be all active affirmative statements, about different topics, such as these:

Jim is washing the dishes.
The boys were eating the apples.
My sister will type your essay.

The following examples are similar because they are all questions :

Is Jim washing the dishes?
Will the apples be eaten by the boys?
Hasn't my sister typed your essay?

These sentences could be grouped together because they are all passive :

The dishes weren't being washed by Jim.
Will the apples be eaten by the boys?
Your essay has been typed by my sister.

The following sentences can go together because they are all declarative, negative, and passive :

The dishes weren't being washed by Jim.
The apples haven't been eaten by the boys.
Your essay won't be typed by my sister.

These sentences, on the other hand, are all active,

<u>affirmative questions</u> :

Will Jim wash the dishes?
Have the boys eaten the apples?
Is my sister typing your essay?

These next example sentences could be put together because they are all <u>active</u> sentences in the <u>past tense</u> :

Jim was washing the dishes.
Weren't the boys eating the apples?
Was my sister typing your essay?

And so on . . . If you have any questions at this point, please ask the experimenter. Press the starting-switch when you are ready to go on.

. . . . .

In the experiment which follows, you will hear the 128 different sentences one at a time, in random order. They can be grouped in many different ways, as we have demonstrated to you; but there is <u>one</u> particular grouping which we want you to try to discover, by trial and error. After each sentence has been presented, say 'yes' or 'no': say 'yes' if you think the sentence belongs to the target-group; say 'no' if you think it does not, or if you don't know. Each time you respond, the experimenter will tell you whether you were correct or wrong. After several presentations, you should have some idea of which kind of sentence is <u>in</u> the target-group, and which kinds are not. Once you have figured out what kind of sentence <u>is</u> in the target-group, you will be able to recognize a target-sentence whenever you hear one, and respond correctly to every presentation thereafter.

During the first few presentations, you will of course have no idea of what kind of sentence to say 'yes' to; therefore, say 'no' to the first few sentences. As long as the experimenter tells you that 'no' is correct, that means you are rejecting non-target sentences correctly. But, when the experimenter says that 'no' is <u>wrong</u>, try to figure out what kind of sentence you just heard, because it will be the first of your target-sentences. After that, say 'yes' or 'no' as you see fit; eventually you will figure out, by trial and error, what kind of sentence to say 'yes' to, and you will say 'no' to all other kinds. After each sentence, the tape-player will stop, so that you may respond; whenever you are ready for the next sentence, press the starting-switch firmly.

Do not try to memorize any sentences. There are 128 of them in all, and they are rather complex. Instead, try to discover a grouping-system based on sentence <u>properties</u>, and keep these in mind, rather than any specific sentences.

If there is anything you do not understand, please

ask the experimenter now; press the starting-switch when you are ready to go on.

. . . . .

The experiment will now begin. After each sentence has been presented say 'yes' if you think it is one of the target-sentences; say 'no' if you think it is not, or if you don't know, such as in the first few presentations. After you have responded, press the starting-switch whenever you are ready for the next sentence."

(Total reading time, as recorded: 6 min., 45 sec.)

APPENDIX E

Complete Analysis of Variance: Yes-Responses as a Function of Sex, Target-feature, and Error-type

Source	SSq	df	MSq	F
Sex (S)	3.955	1	3.955	1.613
Target-features (Tgts):				
Voice (V)	0.564	1	0.564	0.230
Mood (U)	6.799	1	6.799	2.772
Modality (O)	1.221	1	1.221	0.498
V X U	0.861	1	0.861	0.351
V X O	2.127	1	2.127	0.867
U X O	1.033	1	1.033	0.421
V X U X O	0.158	1	0.158	0.065
Sex X Tgts:				
S X V	1.643	1	1.643	0.67
S X U	1.033	1	1.033	0.421
S X O	0.018	1	0.018	0.007
S X V X U	15.470	1	15.470	6.308*
S X V X O	0.018	1	0.018	0.007
S X U X O	0.049	1	0.049	0.02
S X V X U X O	1.643	1	1.643	0.67
*S*s(Sex X Tgts)	117.717	48	2.452	
Error-types (E)	4913.543	7	701.935	156.687**
Sex X E	51.465	7	7.352	1.641
Tgts X E:				
V X E	43.543	7	6.220	1.389
U X E	10.684	7	1.526	0.341
O X E	30.199	7	4.314	0.963
V X U X E	90.232	7	12.890	2.877**
V X O X E	5.404	7	0.772	0.172
U X O X E	56.373	7	8.053	1.798
V X U X O X E	99.311	7	14.187	3.167**
Sex X Tgts X E:				
S X V X E	37.389	7	5.341	1.192
S X U X E	105.998	7	15.143	3.380**
S X O X E	5.951	7	0.850	0.19
S X V X U X E	6.373	7	0.910	0.203
S X V X O X E	32.264	7	4.609	1.029
S X U X O X E	20.482	7	2.926	0.653
S X V X U X O X E	13.322	7	1.903	0.425
*S*s(Sex X Tgts) X E	1505.228	336	4.48	

* $p \leq 0.05$ ** $p \leq 0.01$

B30033